

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Snapshots of proteasomal precursor complexes reveal
chaperone involvement in 20S proteasome biogenesis**

Malte Kock

aus

Bad Oldesloe, Deutschland

2014

Erklärung:

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Frau Dr. Petra Wendler betreut.

Eidesstattliche Versicherung:

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 08.02.2014


(Malte Kock)

Dissertation eingereicht am 11.02.2014

1. Gutachterin: Dr. Petra Wendler

2. Gutachter: Prof. Dr. Roland Beckmann

Mündliche Prüfung am 11.04.2014

Table of contents:

1 Summary/Zusammenfassung	1
2 Introduction	3
2.1 Proteasome function in eukaryotic cells	3
2.2 Structure of the 20S proteasome	4
2.3 Activity of the 20S proteasome	6
2.4 Biogenesis of the 20S proteasome in eukaryotes	7
2.5 Chaperones involved in 20S biogenesis in eukaryotes	9
2.5.1 Pba1 and Pba2	9
2.5.2 Pba3 and Pba4	11
2.5.3 Ump1	12
2.5.4 β subunit pro-peptides and N- or C-terminal extensions	13
2.5.5 Proteasome activators	14
2.6 Electron microscopy and single particle reconstruction	15
2.6.1 Transmission electron microscopy	15
2.6.2 Negative stain EM	16
2.6.3 Cryo-EM	17
2.6.4 CTF correction and filtering	17
2.6.5 Multivariate statistical analysis	18
2.6.6 Angular reconstitution	19
2.6.7 Projection matching	19
3 Materials	20
3.1 Chemicals	20
3.2 Plasmids and strains	20
3.2.1 Plasmids	20
3.2.2 <i>E. coli</i> strains	20
3.2.3 <i>S. cerevisiae</i> strains	20
4 Methods	21
4.1 Cell culture methods	21
4.1.1 Media for <i>S. cerevisiae</i> cultures	21
4.1.2 Medium for <i>E. coli</i> cultures	21
4.1.3 <i>S. cerevisiae</i> cell cultures	22

4.1.4	<i>E. coli</i> cell cultures	22
4.2	Molecular biology methods	22
4.2.1	Plasmid isolation	22
4.2.2	Restriction digestion	22
4.2.3	Transformation of <i>S. cerevisiae</i> cells by electroporation	22
4.2.4	Agarose gel electrophoresis	23
4.3	Protein biochemistry methods	24
4.3.1	Protein precipitation using TCA and NaDOC	24
4.3.2	SDS polyacrylamide gel electrophoresis (PAGE)	24
4.3.3	Coomassie Brilliant Blue stain	25
4.3.4	Amido black stain	25
4.3.5	Silver stain	26
4.3.6	Western blot	26
4.3.7	Cell lysis by bead beater	27
4.3.8	Non-native cell lysis	28
4.3.9	Purification of 15S ^{GFP} complexes for electron microscopy	28
4.3.10	Purification of 15S ^{GFP} complexes for DSS cross-linking experiments	29
4.3.11	Analytical ultracentrifugation	29
4.3.12	Determination of protein concentrations	30
4.3.13	Protein cross-linking using disuccinimidyl-suberate (DSS)	30
4.4	Electron microscopy methods	31
4.4.1	Glow discharging of EM grids	31
4.4.2	Negative stain	31
4.4.3	Preparation of vitrified samples for cryo-EM	32
4.4.4	Nanogold labelling	32
4.4.5	Collection of negative stain data	32
4.4.6	Collection of cryo-EM data	33
4.4.7	Pre-processing of electron microscopy images	33
4.4.8	Multivariate statistical analysis and multi reference alignments	33
4.4.9	Angular reconstitution	34
4.4.10	Refinement	34
4.4.11	Resolution determination	35

4.4.12	Determination of handedness	35
5	Biochemical results	37
5.1	Transformation and expression tests	37
5.2	Purification of proteasome precursor complexes	37
5.3	Titration for DSS H12/D12 cross-linking experiments	39
5.4	Analytical ultracentrifugation of cross-linked proteasome precursor complexes	40
5.5	Proteasome precursor complex cross-linking results	42
6	EM results	46
6.1	EM analysis of purified protein complexes	46
6.2	Classification of EM data	46
6.2.1	Eigenimage analysis	46
6.2.2	Class averages analysis	49
6.3	Refinement	50
6.3.1	15S and 15S ^{GFP} 3D reconstructions	50
6.3.2	20S <i>pre1-1</i> and 20S 3D reconstructions	54
6.3.3	Resolution and dataset statistics	58
6.3.4	Determination of handedness - 15S complex	58
6.3.5	Determination of handedness - 20S <i>pre1-1</i> complex	60
6.4	Subunit fits for negative stain reconstructions - 15S and 15S ^{GFP} complexes ...	60
6.5	Subunit fits for negative stain reconstructions - 20S <i>pre1-1</i> and 20S complexes	65
6.6	Localization of the Pba1-Pba2 heterodimer in the 15S and 20S <i>pre1-1</i> precursor complexes	67
6.7	Localization of Ump1 in the 15S precursor complex	70
7	Discussion	74
7.1	The role of Pba1-Pba2 in 20S biogenesis	74
7.1.1	Alterations of the HbYX motif binding pockets influence Pba1-Pba2 binding modes	74
7.1.2	The position of Pba1-Pba2 in the precursor complexes changes during maturation	75
7.1.3	Possible cross-talk between the α and β ring	76
7.2	The role of Ump1 in 20S biogenesis	77

7.2.1	The predicted structure of Ump1	77
7.2.2	Correlating structure and function of parts of Ump1	78
7.3	An updated model of chaperone-assisted 20S proteasome biogenesis	80
8	References	82
9	Appendix	88
10	Abbreviations	95
11	Acknowledgements	97

1. Summary:

The 26S proteasome is a ~2.5 MDa protein complex essential for protein homeostasis in eukaryotic cells. It consists of the regulatory particle (RP) tasked with the recognition and unfolding of poly-ubiquitinated substrate proteins and a barrel-like 20S core particle (CP) that degrades substrates using its proteolytic activity. Due to its higher complexity compared to functionally similar protein complexes in prokaryotes and archaea, eukaryotic cells employ a number of chaperones to facilitate the correct assembly of the proteasome. This thesis employed biochemical and biophysical techniques to explore the role of the chaperones Pba1, Pba2 and Ump1 in the biogenesis of the 20S proteasome from *S. cerevisiae*. The project was carried out in collaboration with the labs of Prof. Dr. Jürgen Dohmen (University of Cologne) and Dr. Franz Herzog (Gene Center Munich).

While the crystal structure of the Pba1-Pba2 chaperone heterodimer has been solved previously, its exact role in the biogenesis of the 20S proteasome was not entirely clear. This thesis shows that the heterodimer changes its conformation during maturation, initially acting as a scaffold for the assembly of the proteasome α subunits into the α ring, while it mainly blocks the binding of proteasome activators in later stages of maturation.

The third chaperone investigated in this thesis is Ump1, which has so far eluded all attempts to determine the protein structure or its position inside the nascent proteasome. By tracing the amino acid chain of Ump1 inside the 15S proteasome precursor complex, it was determined that Ump1 is monomeric and probably mostly unstructured. The C-terminal part of the protein is found inside the cavity of the proteasome and is involved in binding the chaperone to the proteasome precursor. In addition it may play a part in binding the pro-peptide of the $\beta 5$ subunit, which is essential for proteasome biogenesis. The N-terminal part of Ump1 exits the cavity near the interface where two half-proteasomes dimerize and appears to play a role in this very process. While the exact role of Ump1 in 20S biogenesis could not be determined with absolute certainty, the structural biology data gained from investigating the chaperone should nevertheless prove to be very useful for future studies.

Zusammenfassung:

Das 26S Proteasom ist ein ~2.5 MDa großer Proteinkomplex, welcher für die Protein-Homöostase von eukaryotischen Zellen von essentieller Bedeutung ist. Es besteht aus zwei größeren Subkomplexen: der regulatorische Komplex (RP) erkennt und entfaltet poly-ubiquitinierte Substrat-Proteine, während im Kern des Komplexes die Proteine durch die proteolytische Aktivität des 20S Proteasoms gespalten werden. Da das Proteasom eukaryotischer Zellen einen komplexeren Aufbau besitzt als funktionell ähnliche Proteinkomplexe in Prokaryoten und Archaeen, sind für seinen korrekten Zusammenbau mehrere Chaperone nötig. In dieser Arbeit wurden verschiedene biochemische und biophysikalische Methoden dazu verwendet, die Rolle der Chaperone Pba1, Pba2 und Ump1 in der Biogenese der 20S Proteasoms aus *S. cerevisiae* aufzuklären. Das Projekt wurde in Zusammenarbeit mit den Arbeitsgruppen von Prof. Dr. Jürgen Dohmen (Universität Köln) und Dr. Franz Herzog (Genzentrum München) durchgeführt.

Obwohl die Kristallstruktur des Pba1-Pba2 Heterodimers bereits gelöst wurde, war die genaue Rolle der beiden Chaperone in der Biogenese des 20S Proteasoms bisher nicht restlos geklärt. In dieser Arbeit konnte gezeigt werden, dass die Chaperone während der Biogenese ihre Konformation im proteasomalen Vorläuferkomplex ändern. Während sie in der Frühphase der Biogenese als eine Art Gerüst für den Zusammenbau des proteasomalen α -Rings fungieren, verhindern sie im weiteren Verlauf hauptsächlich die Bindung von Aktivatorkomplexen an das unfertige Proteasom.

Das dritte Chaperon welches in dieser Arbeit untersucht wurde ist Ump1. Für dieses Protein konnte bisher keine Struktur ermittelt werden und seine genaue Position und Stöchiometrie in proteasomalen Vorläuferkomplexen war unbekannt. Durch Lokalisation unterschiedlicher Teile der Aminosäurekette des Proteins im 15S Vorläuferkomplex konnte ermittelt werden, dass das Protein als Monomer vorliegt und wahrscheinlich größtenteils unstrukturiert ist. Der C-terminale Bereich des Proteins befindet sich auf der Innenseite des Komplexes und ist für die Bindung des Chaperons an den Komplex verantwortlich. Dieser Bereich spielt möglicherweise eine Rolle bei der Bindung des Propeptids der $\beta 5$ Untereinheit, welches für die Biogenese des 20S Proteasoms essentiell ist. Der N-terminale Part von Ump1 wurde nahe der Kontaktfläche lokalisiert, über die zwei Halb-Proteasomen dimerisieren und spielt aller Wahrscheinlichkeit nach eine Rolle in diesem Prozess. Obwohl die genaue Aufgabe von Ump1 in der Biogenese des 20S Proteasoms nicht mit absoluter Sicherheit geklärt werden konnte, so liefern die strukturellen Daten doch einen wertvollen Beitrag für zukünftige Studien.

2. Introduction:

2.1 Proteasome function in eukaryotic cells:

The 20S core particle is part of the 26S proteasome, a ~2.5 MDa protein complex tasked with degradation of poly-ubiquitinated proteins in eukaryotic cells. Each part of this complex plays a vital role, resulting in specific and highly effective cleavage of substrates. While the structure of the 20S core particle has long been known, recent years have shed light on its interplay with the ATPase and non-ATPase subunits of the 19S regulatory particle. The non-ATPase subunits (also called the lid) can recognize ubiquitinated substrates and hand them over to an ATPase ring, which unfolds them by converting the chemical energy of ATP hydrolysis into mechanical force. While the exact mechanism of unfolding still remains elusive, it is known that the unfolded substrate is channeled through the pore of the ATPase ring into the 20S core particle, where the substrate is proteolytically cleaved into smaller peptides.

Proper function of the proteasome is vital to all eukaryotic life. In addition to its obvious role in protein homeostasis, the proteasome also confers resistance to elevated temperatures as well as oxidative stress (Friant et al. 2003; Shang & Taylor 2011; Bader & Grune 2006). It plays an important role in the cell cycle, degrading cyclins and transcription factors to drive the cycle forward (Chesnel et al. 2006; Havens et al. 2006). Malfunction of proteasomes has been implicated in a wide array of phenotypes including a higher rate of DNA damage (Jacquemont & Taniguchi 2007), mitochondrial dysfunction (Taylor & Rutter 2011) and accumulation of misfolded or otherwise damaged proteins (Takalo et al. 2013). This has made the proteasome a prime target for fighting diseases in humans, ranging from neurodegenerative diseases to cancer and tumors. Bortezomib was the first substance approved for use in humans and it works through inhibition of the catalytic activity of the 20S proteasome, disrupting the protein homeostasis of cancer cells.

Prokaryotes and archaea possess much simpler proteasome analogues compared to eukaryotic cells (Bochtler et al. 1999). In prokaryotes the proteins HslU and HslV can assemble into complexes reminiscent of the 26S proteasome, with parts being responsible for substrate recognition and translocation while others handle the degradation. These complexes are however less important for cell survival than their counterpart in eukaryotes because protein degradation is a much more distributed process in prokaryotes. Proteasomes in archaea usually consist of complexes made from 2 individual subunits reminiscent of α and β subunits found in the 20S core particle of eukaryotes.

2.2 Structure of the 20S proteasome:

In eukaryotes the 20S proteasome is made up of 28 subunits in total, 14 of which are unique. Seven α and seven β subunits assemble into rings, giving the proteasome a characteristic $\alpha_7\beta_7\beta_7\alpha_7$ composition (Groll et al. 2000). Figure 1A and 1B show the 3D representation of a 20S proteasome from *S. cerevisiae*, revealing a multi chambered barrel-like structure. The overall size and molecular weight of 20S proteasomes are more or less conserved among eukaryotes, being 150 Å in height, 120 Å in diameter and around 750 kDa.

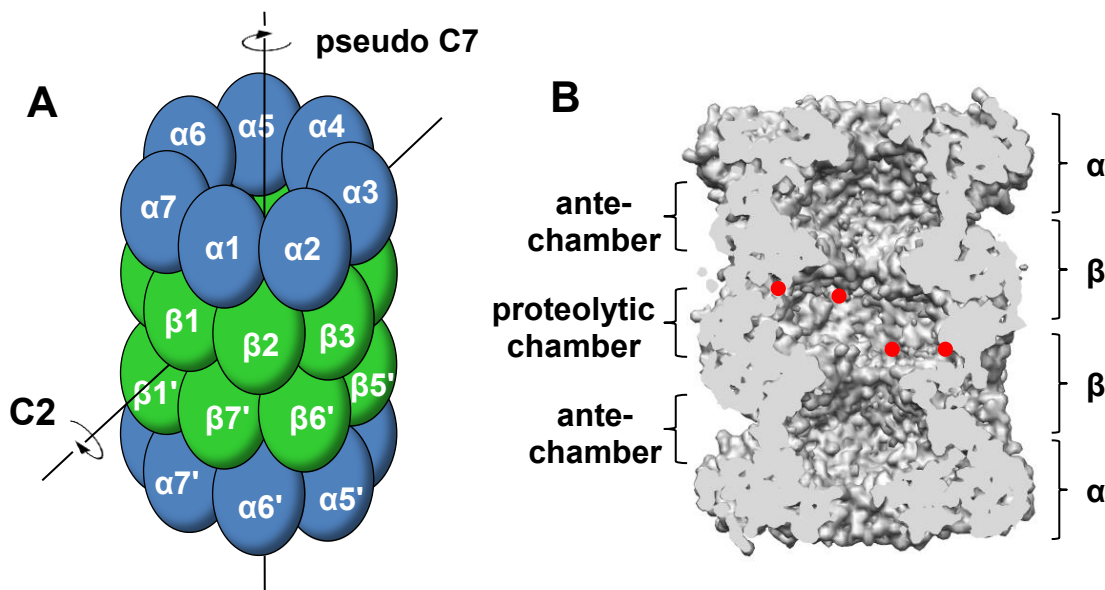


Figure 1: (A) Schematic representation of 20S proteasome subunit arrangement in the mature complex. Eukaryotic proteasome possess a pseudo C7 symmetry axis running along the pore and a real C2 symmetry axis separating the two halves between the subunits β_1/β_1' and β_4/β_4' . (B) Cut-open side view of the 20S proteasome showing the antechambers and the proteolytic chamber. Approximate locations of proteolytic sites have been marked in red. The volume is based on the crystal structure of the *S. cerevisiae* 20S proteasome (PDB: 1G0U (Groll et al. 2000)) filtered to 5 Å resolution.

The α subunits control access to the antechambers by rearranging their N-terminal regions, effectively opening or closing the pore (Groll et al. 2000). All α subunits share a high degree of similarity with regard to their 3D structure (figure 2A), although their amino acid sequences are only poorly conserved (21-35 % sequence similarity, (Groll 2004)). Each subunit consists of two 5-stranded anti parallel β -sheets, sandwiched between helix H1-2 on one side and H3-5 on the other. N-terminal extensions reaching into the pore contain an additional helix H0.

The β subunits are situated between the α rings at the middle of the barrel, forming the proteolytic chamber of the 20S. Their 3D structure is highly similar to the α subunits (figure 2B), showing the same arrangement of 2 anti-parallel β -sheets flanked by α helices, although

there is an even lower amino acid conservation among the subunits (11-26 % sequence similarity, (Groll 2004)). They also don't have N-terminal H0 helices. Instead several β subunits have other appendages important for proteasome structure and function. $\beta 2$'s C-terminal extension wraps around the neighboring subunit $\beta 3$. The subunit $\beta 7$ also has a long C-terminal tail that intercalates between $\beta 1$ and $\beta 2$ of the neighboring β -ring.

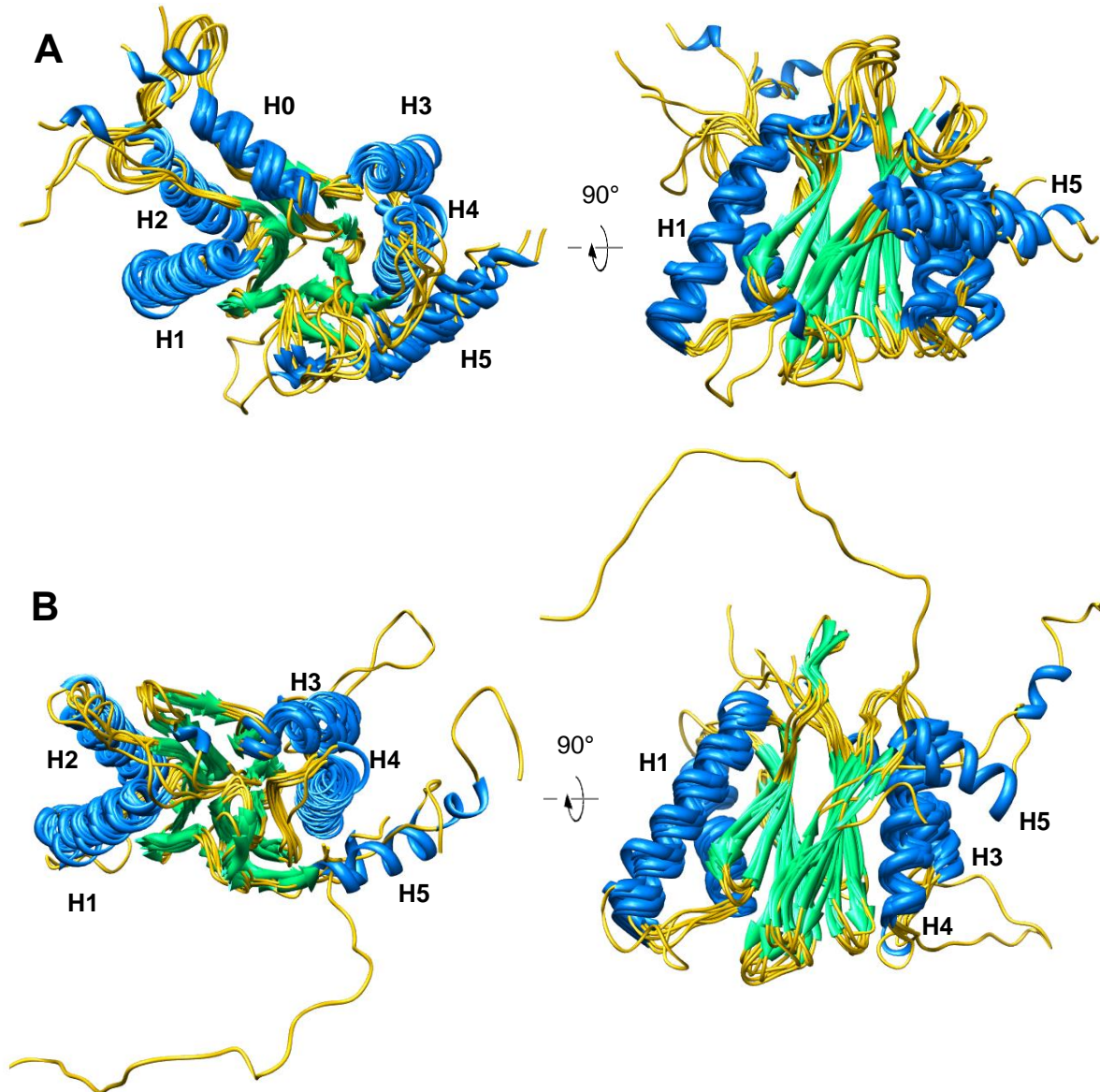


Figure 2: Overlays of all proteasome α subunits (A) and β subunits (B) from *S. cerevisiae* showing the conservation of the 3-dimensional structure. Proteins are depicted as ribbons, with helices shown in blue, sheets in light green and loops in yellow. Helices are denoted H0-H5 starting from the N- to the C-terminus. Atomic models were extracted from the crystal structure of the 20S proteasome (PDB: 1G0U (Groll et al. 2000)).

Due to the high degree of similarity between α and β subunits and their ring-like arrangements the proteasome features several symmetry axes. One runs along the middle of the pore and

represents a pseudo rotational 7-fold symmetry while the other is situated between two half-proteasomes, giving the whole assembly a centrosymmetric 2-fold symmetry (figure 1A).

2.3 Activity of the 20S proteasome:

In eukaryotes the activity of the 20S proteasome usually refers to the cleavage of substrates in the proteolytic chamber, although opening and closing of the α ring pore are also important events with respect to substrate degradation. These events are triggered by the presence or absence of regulatory factors, like the ATPases of the 19S regulatory particle (RP). These regulatory proteins can interact with 5 canonical lysines (6 in mammals) in shallow pockets between subunits $\alpha 1/\alpha 7$, $\alpha 2/\alpha 3$, $\alpha 3/\alpha 4$, $\alpha 4/\alpha 5$ and $\alpha 5/\alpha 6$ using their C-terminal HbYX motifs (i.e. a hydrophobic amino acid, followed by tyrosine, followed by a single random amino acid). Binding events to the pocket trigger a rearrangements of α subunit N-terminal pore loops, opening the pore for potential substrates carried by the regulatory particle. This has been likened to a “key-in-a-lock” mechanism (Smith et al. 2008). In yeast, it has been shown that mutation of the HbYX tyrosine residue can abolish gating activity while keeping the interaction between 20S and RP intact, emphasizing the importance of proper control over the pore opening mechanism (Saeki & Tanaka 2007). Even though the binding of regulatory proteins widens the pore, it is still narrow in its open form (~ 13 Å). This suggests that proteins have to be unfolded before entering the 20S. Curiously there have been examples in which partial degradation of unstructured protein regions drives activation of transcription factors, like in the NF κ B-like Spt23p and Mga2p from *S. cerevisiae* (Rape & Jentsch 2002). In these cases structured parts of the target protein stay outside the pore, forgoing complete degradation in favor of activation.

Although the β ring consists of 7 individual subunits, only 3 of them ($\beta 1$, $\beta 2$ and $\beta 5$) are proteolytically active. These subunits are synthesized with N-terminal pro-peptides that have to be cleaved off first for them to become active. Apart from their role in assembly of the β ring these pro-peptides also protect the active site threonine (Thr1) from acetylation prior to 20S assembly (Jäger et al. 1999). Deprotonation of the Thr1 hydroxyl group enables this residue to perform a nucleophilic attack on potential substrates, while a charge relay system of conserved downstream residues (Asp17, Lys33, Ser129, Asp166 and Ser169) are thought to facilitate this reaction by displacing the Thr1 proton (Heinemeyer et al. 2004). Although Thr1 is a conserved residue in $\beta 1$, $\beta 2$ and $\beta 5$, slight variations of amino acid residues in the active center of the different β subunits give rise to substrate specificity. An overview of catalytic β subunits and their activities is given in table 1. Cleavage of substrates usually leaves small

peptides 7-9 amino acids in length, although this can vary from one substrate to another (Voges et al. 1999).

Table 1: Catalytic activities of 20S proteasome β subunits. Primary activities are shown in *italic*, secondary activities in plain letters. PGPH = peptidyl-glutamyl peptide-hydrolyzing (cleavage after acidic residue). ChyT-l = chymotrypsin-like (cleavage after hydrophobic residue). T-l = trypsin-like (cleavage after basic residue). (Orlowski & Wilk 2000; Heinemeyer et al. 1997; Dick et al. 1998)

Catalytic subunit	Proteolytic activity
$\beta 1$	<i>PGPH</i> , ChyT-l
$\beta 2$	<i>T-l</i> , ChyT-l
$\beta 5$	<i>ChyT-l</i>
$\beta 1i$	<i>ChyT-l</i>
$\beta 2i$	<i>T-l</i>
$\beta 5i$	<i>ChyT-l</i>
$\beta 5t$	<i>ChyT-l</i>

In higher vertebrates, interferon γ can induce the expression of the catalytic subunits $\beta 1i$, $\beta 2i$ and $\beta 5i$, which are assembled into the proteasome in place of their wild type counterparts (Wang & Maldonado 2006). These subunits confer altered cleavage specificity to the proteasome that produces peptides better suited for presentation by major histocompatibility complex I molecules on the cell surface. An altered proteasome containing these subunits is referred to as the immunoproteasome. Similarly the cortical epithelial cells of the thymus express a third type of proteasome which is almost identical to the immunoproteasome, but $\beta 5i$ has been replaced by $\beta 5t$. This subunit displays a markedly reduced chymotrypsin-like activity compared to its wild type counterpart. While the function of the epithelial cells in T-lymphocyte selection has been established, the exact role of the thymoproteasome in this process is not well understood (Murata et al. 2007).

2.4 Biogenesis of the 20S proteasome in eukaryotes:

Biochemical and structural biology data from eukaryotes indicates that the assembly of 20S is not a spontaneous or autonomous process, but instead highly regulated and driven by chaperones (figure 3). The biogenesis of 20S is thought to begin with the formation of α rings, placing each α subunit into a specific position in the ring. In *S. cerevisiae* this task is facilitated by the chaperones Pba1-4 (**P**roteasome **b**iogenesis **a**ssociated) based on findings in mammalian cells concerning their orthologues PAC1-4 (**P**roteasome **A**ssembling **C**haperone) (Hirano et al. 2005; Hirano et al. 2006; Kusmierczyk et al. 2008). The complete α ring serves

as a docking platform for the β subunits. $\beta 2$ is the first subunit to join the α ring, followed by $\beta 3$ and $\beta 4$. The heterodimer Pba3-Pba4 is lost in the process. In human cell the incorporation of $\beta 2$ requires another chaperone called hUmp1 (**h**uman **U**biquitin-**m**ediated **p**roteolysis) (Fricke et al. 2007). In yeast the ortholog Ump1 follows $\beta 3$ and $\beta 4$ into the complex, though it is unclear at which point exactly and where it binds (Ramos & Dohmen 2008). After Ump1 the subunits $\beta 5$, $\beta 6$ and $\beta 1$ are incorporated into the ring. This intermediate is referred to as the 15S or half-mer ($-\beta 7$). At this stage one of the proteasome activators like PA200 or PA700 can bind to the 15S. In *S. cerevisiae* this is usually done by Blm10, though the 19S regulatory complex can apparently substitute for Blm10 to some extent (Marques et al. 2007). The addition of the last subunit $\beta 7$ leads to rapid dimerization of 2 half-proteasomes through intercalation of the C-terminal extension between $\beta 1$ and $\beta 2$ in the opposing β ring (Li et al. 2007). Once the half-proteasomes are joined the catalytic β subunits can activate their counterparts in the opposing ring by cleaving off their pro-peptides in a process called auto-activation (Schmidtke et al. 1996). The pro-peptides of the inactive subunits $\beta 6$ and $\beta 7$, as well as the chaperones Ump1, Pba1 and Pba2 are also processed by active β subunits. A proteasome mutant called 20S *pre1-1* (Heinemeyer et al. 1993) can be used to mimic the penultimate step in 20S biogenesis, when the core particle is fully assembled but chaperones are still attached. This mutant carries an S142F mutation in $\beta 4$ and retains less than 5 % of chymotrypsin-like activity that is needed to process the $\beta 5$ pro-peptide as well as the chaperones. The β subunit activation and disposal of the chaperones concludes the 20S biogenesis.

The initial steps of 20S proteasome biogenesis take place in the cytoplasm. Experiments with yeast cells carrying a truncated version of the nuclear import protein karyopherin α showed that 13-16S proteasomal precursor complexes accumulate in the cytoplasm, suggesting that under wild type conditions these precursors are imported into the nucleus where the remaining biogenesis steps take place (Lehmann et al. 2002). Further evidence is found in α subunits like $\alpha 4$ which carry nuclear import signal sequences, while β subunits are devoid of them. This points to the fact that the β subunits need to be bound to α rings in order to be imported into the nucleus. Indeed in yeast the inner surface of nuclear envelopes is dotted with proteasomes, comprising up to 80 % of all the proteasomes in the cell (Enenkel et al. 1998).

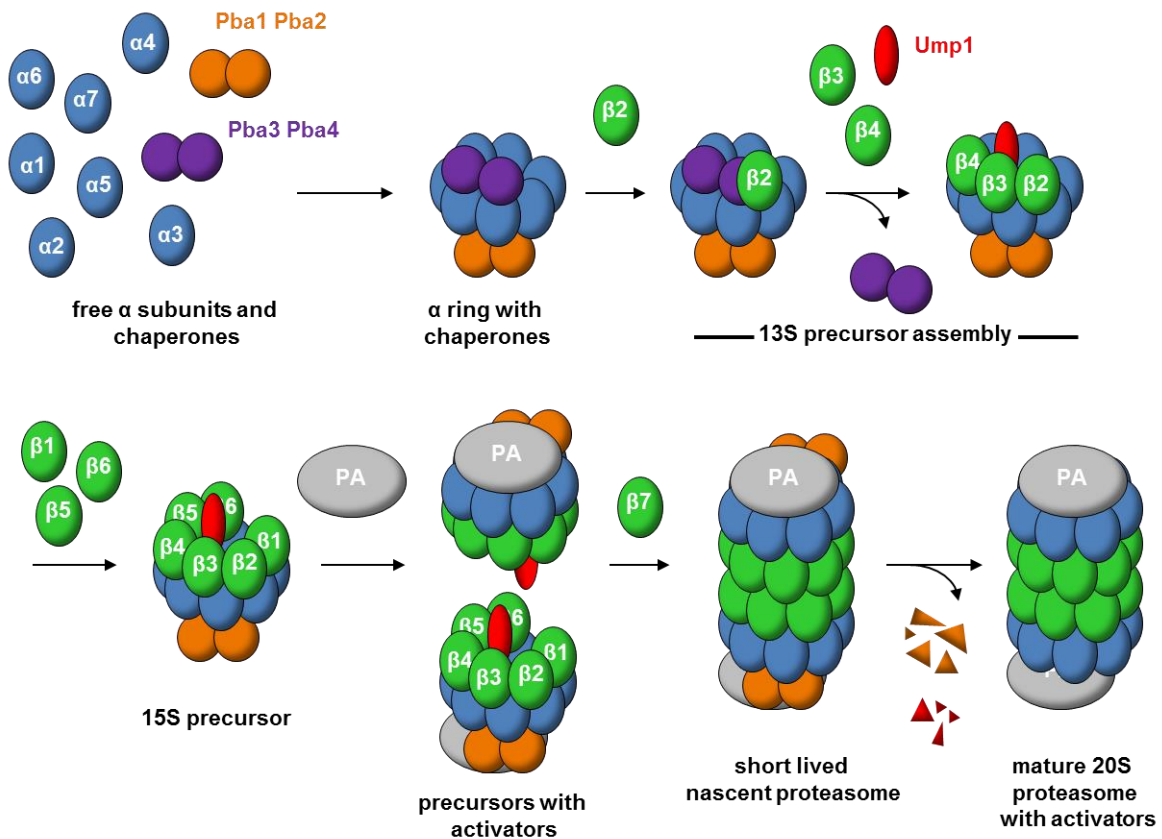


Figure 3: Biogenesis pathway showing the different subunits and chaperones involved in the assembly of fully functional 20S proteasomes in *S. cerevisiae* (based on Ramos & Dohmen 2008). In the first step the α ring is assembled in the correct order by the chaperone heterodimers Pba1-Pba2 and Pba3-Pba4. The α ring serves as a scaffold for the first β subunits $\beta 2$ -4. At this stage the incomplete proteasome is joined by the chaperone Ump1, followed by the β subunits $\beta 5$, $\beta 6$ and $\beta 1$, forming the 15S precursor complex. The 15S is bound by a proteasome activator (PA) and the final subunit $\beta 7$, which leads to rapid dimerization of two half-proteasomes. The proteolytic β subunits $\beta 1$, $\beta 2$ and $\beta 5$ are activated by their counterparts in the opposing ring and they digest the chaperones Pba1-Pba2 as well as Ump1, leading to the formation of the mature 20S proteasome.

2.5 Chaperones involved in 20S biogenesis in eukaryotes:

2.5.1 Pba1 and Pba2:

In eukaryotic cells the subunit $\alpha 7$ can self-assemble into homo-heptameric rings, which are also able to dimerize into two-tiered ring-like structures (Gerards et al. 1997). This suggests an assembly mechanism in which additional factors keep this from happening by ensuring that every subunit is only present once and in the correct order. The first evidence that assembly of the α ring is a chaperone mediated process was found in human cell with the identification of PAC1 and PAC2 (Hirano et al. 2005). Both proteins were found to be associated with

proteasomal precursors containing α and β subunits. Their yeast orthologues were termed Pba1/Poc1 and Pba2/Poc2, but they only share weak sequence similarity with their mammalian counterparts. PAC1 and PAC2 are usually isolated in stoichiometric amounts, suggesting they form heterodimers. They associate directly with $\alpha 5$ and $\alpha 7$, suggesting a role in forming α rings in the correct order. Both hypotheses were confirmed with the publication of a crystal structure containing mature 20S from *S. cerevisiae*, reconstituted with recombinantly expressed Pba1 and Pba2 (Stadtmueller et al. 2012). This structure places the chaperone heterodimer on the outer surface of the proteasome, where it interacts with pockets between the α subunits in a manner reminiscent of proteasome activators (i.e. using HbYX motifs on their C-terminal ends). In this structure Pba2 interacts with an inter-subunit pocket between $\alpha 6$ and $\alpha 7$ that doesn't seem to be used by other activators (Saeki & Tanaka 2007). While it is evident that Pba1-Pba2 could steer $\alpha 5$ -7 into the right position through direct interactions, there are no contacts to subunits $\alpha 1$ -4 resolved in the crystal structure.

Both chaperones share a lot of similarities concerning their 3D structure. They consist of a 4-stranded parallel β sheet, extended by antiparallel sheets and flanked by 2 helices on either side. The main differences between both proteins arise from loop regions connecting the structured parts and their C-terminal regions. The latter is rather short in Pba1 and consists of a helical region that interacts with $\alpha 5$. In Pba2 this region is much longer and more unstructured, consisting of several loops and 2 short helices. The terminal helix is used to interact with $\alpha 7$.

Deletion or depletion of Pba1-Pba2 or one of its respective mammalian orthologues can have different effects, depending on the species. In human cells the siRNA mediated knockdown of PAC1-PAC2 resulted in the formation of aberrant proteasome precursors, containing all α subunits in the aforementioned two-tiered ring-like structures (Hirano et al. 2005). This leads to a severe drop in intracellular proteasome activity through the formation of dead end intermediates that can no longer progress down the proteasome biogenesis pathway. In yeast on the other hand the deletion of Pba1-Pba2 has next to no effect on cell viability, as long as the proteasome transcription factor Rpn4 is present (Le Tallec et al. 2007). This suggests that deletion of Pba1-Pba2 leads to the formation of partially defective proteasomes, but the cells can compensate by up-regulation of proteasome biosynthesis. Deletion mutants therefore accumulate more proteasome precursor complexes.

While it is clear that Pba1 and Pba2 enter the proteasome biogenesis pathway right at the beginning by helping in the assembly of the α ring, it is unclear when they leave the nascent proteasome. The orthologues PAC1 and PAC2 were found to have a half-life approximately

equal to the time it takes to assemble 20S, hence it was assumed they are degraded by the nascent proteasome (Hirano et al. 2005). Using a proteasome inhibitor called MG132 that keeps the $\beta 5$ subunit from becoming active results in mature 20S associated with PAC1-PAC2 heterodimers. This suggests that Pba1 and Pba2 or their orthologues stay almost all the way until the end of the biogenesis pathway, until they are degraded. This would however put them at odds with proteasome activators using the same binding pockets between α subunits. More importantly, degradation by the nascent proteasome would also require at least a partial unfolding of the chaperones and translocation into the proteolytic chamber, none of which can be accomplished by 20S. Due to this conflicting information, the possibility remains that Pba1-Pba2 simply dissociate at the end of 20S biogenesis, after which they are reused by the cells just like Pba3-Pba4 (see below).

2.5.2 Pba3 and Pba4:

When investigating early proteasome precursors containing PAC1 and PAC2, two more chaperones were identified that are associated with the α ring. They were named PAC3 and PAC4 in human cells, while yeast orthologues are called Pba3/Poc3/Dmp2 and Pba4/Poc4/Dmp1. Although they are associated with complexes containing a complete α ring as well as $\beta 2$, they are absent from later intermediates containing hUmp1 (Hirano et al. 2006), indicating that the binding of these chaperones is mutually exclusive. Like Pba1-Pba2, Pba3 and Pba4 form heterodimers. A crystal structure of Pba3-Pba4 from *S. cerevisiae* places these chaperones in direct contact with $\alpha 5$, however in contrast to Pba1-Pba2 they are located inside the ring, where they would face $\beta 4$ in the mature 20S (Yashiroda et al. 2008). This also explains why these chaperones are lost when the β ring assembles and why (unlike the other chaperones involved in 20S biogenesis) they are not degraded by the nascent proteasome. Its position in the α ring will put Pba3-Pba4 in contact with 2 more subunits, $\alpha 4$ and $\alpha 6$. It is therefore thought that Pba3-Pba4 prevents $\alpha 3$ from binding directly to $\alpha 5$, ensuring the incorporation of $\alpha 4$ into the ring. This means Pba3-Pba4 could fulfil a similar role as Pba1-Pba2, helping with the correct assembly of the α ring. However both heterodimers cannot substitute for one another (Hirano et al. 2006).

Like Pba1-Pba2, the chaperones Pba3 and Pba4 also share a large degree of structural similarity. Both consist of a 6-stranded anti parallel β sheet packed against 2 helices. In the heterodimer the sheets form a β sandwich with the helices situated on either side.

Since Pba3 and Pba4 work only in tandem, deletions and knockdowns of one protein also affect the other. Depletion of PAC3 in human cells by siRNA mediated knockdown lead to

accumulation of ubiquitinated proteins in cells, indicating severe proteasome malfunctions (Hirano et al. 2006). It also causes an accumulation of free α subunits as well as PAC1-PAC2 complexes, meaning that PAC3-PAC4 probably have a role in nucleation of the α ring assembly and bind well before PAC1-PAC2. Deletion of Pba3-Pba4 from *S. cerevisiae* caused the assembly of α rings lacking subunit α 4 (Yashiroda et al. 2008). However when the non-essential subunit α 3 was deleted as well, it caused the incorporation of two α 4 subunits instead (Kusmierczyk et al. 2008). These conflicting results warrant further research into the relevance of Pba3-Pba4 for proteasome biogenesis.

2.5.3 Ump1:

Ump1 was first found to play a role in ubiquitin-mediated proteolysis, hence the name (Ramos et al. 1998). Of all the chaperones involved in 20S biogenesis it has gotten the most attention in science, although information is often conflicting between species. These discrepancies likely arise from the fact that mammalian cells can assemble proteasomes with β subunits not present in yeast (e.g. β 1i, β 2i and β 5i from immunoproteasomes) and therefore require slightly altered functions of Ump1. In yeast, Ump1 binds to the 13S precursor after β 2, β 3 and β 4 have been incorporated into the ring (Li et al. 2007). In human cells on the other hand the ortholog hUmp1 can bind to the α ring independently of β subunits *in-vitro* and *in-vivo* (Fricke et al. 2007). It is therefore thought to enter the biogenesis pathway much sooner than its yeast counterpart, paving the way for correct incorporation of the first β subunit β 2. Incorporation of the chaperone into the precursor seems to have great significance for all the downstream assembly processes, as knockdown of hUmp1 impairs the recruitment of β 5 to the ring. After dimerization of two half-proteasomes, Ump1 is degraded by the nascent proteasome. Mutant proteasomes deficient in proteolytic activity still carry Ump1, but it is not accessible to polyclonal antibodies, suggesting that it has been encased in the newly formed proteasome (Ramos et al. 1998). In yeast the deletion of Ump1 is not lethal, but makes cells hypersensitive to various forms of stress (Ramos et al. 1998; Chen et al. 2006). These cells accumulate proteasome precursors and 20S with unprocessed β subunits, suggesting severe defects in proteasome assembly. This also insinuates that Ump1 has a role in the processing of β subunit pro-peptides. In contrast to yeast, human cells with knocked down hUmp1 do not show an accumulation of 20S with unprocessed β subunits. This is probably due to the fact that the chaperone is needed to incorporate the first β subunit into the nascent proteasome, thus the biogenesis does not move forward.

Apart from its direct role in proteasome biogenesis Ump1 and its orthologues have also been implicated in the trafficking of precursor complexes. hUmp1 seems to target incomplete proteasomes to the endoplasmatic reticulum, where most of the proteasome assembly takes place in mammalian cells (Fricke et al. 2007). In yeast on the other hand they were targeted to the inner side of the nuclear envelope which connects to the endoplasmatic reticulum (Lehmann et al. 2002).

2.5.4 β subunit pro-peptides and N- or C-terminal extensions:

While they cannot be strictly classified as independent chaperones, the pro-peptides and other extensions of β subunits nevertheless have a great effect on the assembly of the 20S proteasome. Perhaps the biggest influence is seen in the unusually long $\beta 5$ pro. Not only has deletion of this sequence been proven to be fatal in yeast (Chen & Hochstrasser 1996), there is also a large degree of interplay with other proteasome subunits. The lethality of this deletion can be avoided through an additional deletion of Ump1, emphasizing the importance of interactions between these factors (Ramos et al. 1998). An overexpression of $\beta 7$ was also able to rescue cells with a $\beta 5$ pro deletion (Li et al. 2007). In this case the C-terminal extension (CTE) of $\beta 7$ was vital to create this effect. During dimerization of two half-proteasomes the $\beta 7$ CTE intercalates between subunits $\beta 1$ and $\beta 2$ of the opposing ring, stabilizing the nascent proteasome while also helping in the activation of $\beta 1$ (Ramos et al. 2004). While not fatal, a deletion of the $\beta 7$ CTE lead to an accumulation of precursor complexes inside the cell, as this subunit was no longer incorporated into β rings. A deletion of $\beta 6$ pro was also able to rescue cells carrying $\beta 5\Delta$ pro (Li et al. 2007). Just like the latter, deletion of the $\beta 6$ N-terminal extension (NTE) was fatal in yeast, but could be rescued by Δ ump1. Overall this data suggests a high degree of interplay between the $\beta 5/\beta 6$ pro-peptides and Ump1, as well as partial redundancy of $\beta 5$ pro and $\beta 7$ pro for half-proteasome dimerization. While the $\beta 6$ NTE seems to promote half-proteasome dimerization, the presence of $\beta 6$ pro has an inhibitory effect. On the opposite side of the ring the subunit $\beta 2$ features both a CTE as well as an N-terminal pro-peptide. The CTE wraps around the neighbouring subunit $\beta 3$. This extension seems to be of paramount importance to the incorporation of $\beta 3$ into the ring, as a deletion proves fatal in yeast (Ramos et al. 2004). While deletion of the N-terminal $\beta 2$ pro does not have the lethality of $\beta 5\Delta$ pro, it nevertheless impairs proteasome assembly to some degree and makes cells hypersensitive to elevated temperatures (Jäger et al. 1999). In contrast, $\beta 1$ pro is completely dispensable, as the premature acetylation of $\beta 1$ Thr1 and loss of PGPH activity have little effect on cell viability (Heinemeyer et al. 1997). All phenotypes caused by $\beta 1\Delta$ pro, $\beta 2\Delta$ pro

and $\beta 5\Delta\text{pro}$ could be mitigated to some extent by supplying these pro-peptides *in trans*. While catalytic activity is still well below wild type level in $\beta 2\Delta\text{pro}$ and $\beta 5\Delta\text{pro}$ (and completely abolished in $\beta 1\Delta\text{pro}$), proteasome assembly is largely restored in the first two examples. This points to $\beta 2\text{pro}$ and $\beta 5\text{pro}$ having dual roles in proteasome assembly and folding of their active centres, while $\beta 1\text{pro}$ solemnly protects the active centre Thr1 from acetylation.

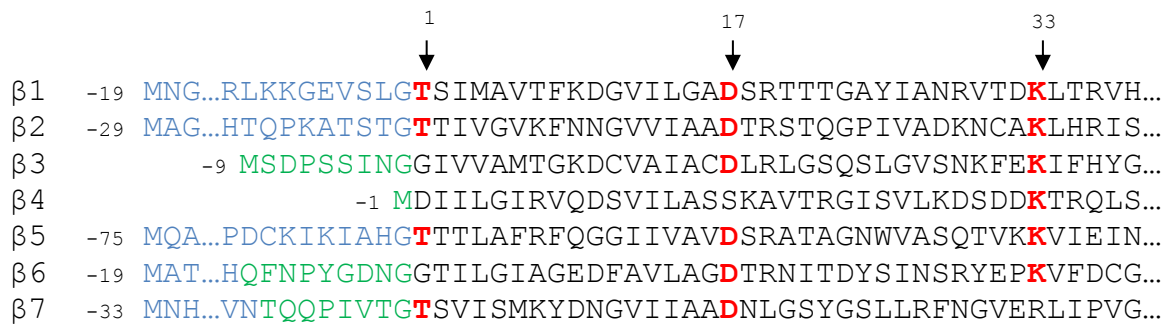


Figure 4: N-terminal amino acid sequences of proteasome β subunits from *S. cerevisiae* showing pro-peptides (light blue), N-terminal extensions (green) and residues important for the catalytic activity, like Thr1, Asp17 and Lys33 (red). Due to the large variation in pro-peptide length, some sequences have been shortened. The starting residues of these sequences are indicated by numbers in front of the rows.

2.5.5 Proteasome activators:

Although it has been established that proteasome activators like PA200/Blm10 and PA700/19S regulatory particles can bind proteasome precursors and stabilize them, their exact role in this process is poorly understood. A deletion of Blm10 has only modest effects on cell viability, possibly because the 19S regulatory particle has partially redundant functions in precursor stabilization. Consequently a mutation in the regulatory protein Rpn2 that prevents attachment of 19S to the α ring impairs 20S proteasome formation when coupled with Δblm10 (Marques et al. 2007). A similar synthetic inhibitory effect is observed when Δblm10 is combined with a deletion of the $\beta 7$ CTE. While this would suggest that Blm10 is a promoter of proteasome maturation, there is also conflicting information in which Δblm10 causes faster β subunit processing and Ump1 degradation, meaning Blm10 could fulfil an inhibitory role (Fehlker et al. 2003). Further discrepancies arise from the fact that proteasome activators use the same binding sites as the chaperones Pba1 and Pba2, which should in principle make their occurrence in precursor complexes mutually exclusive.

2.6 Electron microscopy and single particle reconstruction

2.6.1 Transmission electron microscopy:

Transmission electron microscopy (TEM) refers to the direct visualization of small objects ranging in size from cells down to single proteins by use of an electron microscope. The schematic representation of an electron microscope is shown in figure 5. Inside a vacuum column, an electron source accelerates electrons toward an assortment of condenser lenses and apertures before they hit the specimen. The acceleration voltage is usually in the range of 80-300 kV, resulting in an electron wavelength ranging from 0.155 Å to 0.041 Å. Lenses use magnetic fields to manipulate and focus the electron beam, while apertures act as a filter for electrons scattered at very high angles. The specimen can elastically or inelastically scatter some of the incident electrons, while others are backscattered, absorbed by the sample or simply pass through without any kind of interaction. This requires the use of fixatives to protect the sample from beam damage. Often this is achieved by embedding samples in a heavy metal stain, but the proteins can also be cryogenically frozen (see below). The elastically scattered beam carries all the spacial information needed to reconstruct an image. For visualization the beam is magnified and projected onto a fluorescent screen by the objective and projection lenses. Permanent recording of frames requires the use of film, a CCD camera or a direct electron detector in place of the fluorescent screen. In structural biology this is widely exploited to gain high resolution information of protein complexes. In contrast to crystallographic methods, phases are preserved throughout the imaging process, so that the crystallographic phase problem is averted. TEM also does not require the sometimes artificial packing of proteins into a crystal lattice, but can be applied to protein complexes in near-native states. Consequently it does allow a slightly higher degree of heterogeneity in the sample, as protein complexes are analysed on a per-particle basis. Shortcomings of TEM compared to crystallographic methods include a severely decreased resolution (for most samples). Although the short wavelength of the electron beam of a transmission electron microscope suggests the potential to obtain high resolution information well beyond 0.5 Å, the real resolution is limited by the quality of the electron source, spherical aberration of lenses, resolution of the recording device, sample quality and post-TEM processing of data.

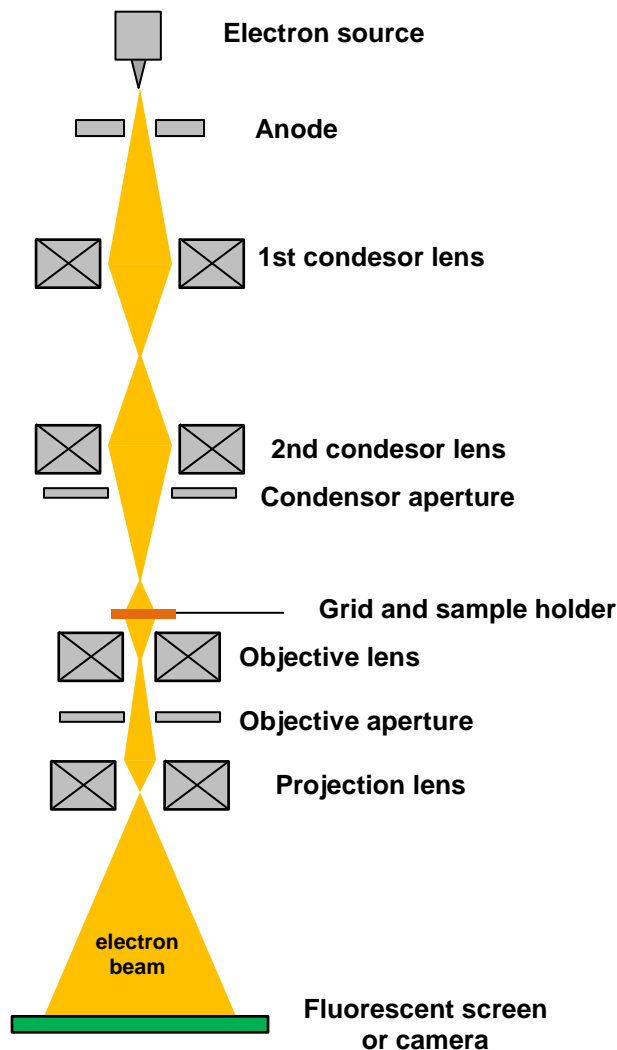


Figure 5: Schematic drawing of a transmission electron microscope, showing the different lenses and apertures necessary to visualize specimen at high resolution. Electrons are extracted from an electron source and accelerated toward the anode, before passing a series of condenser lenses. The latter are responsible for forming a coherent beam from electrons emitted by the electron source. After passing through or interacting with the sample, the beam is again focussed by an objective lens, before being spread by the projections lens for imaging on the fluorescent screen or camera.

2.6.2 Negative stain EM:

In negative stain EM the sample is embedded in a heavy metal stain to protect it from beam damage and to increase the contrast by use of high electron density material. It works on both single proteins and thin sections of cells, providing a wide array of possible uses and applications. The name refers to the fact that the surrounding area of the specimen is stained, not the specimen itself (in contrast to positive staining). For preparation the sample is briefly mixed with the staining agent, before excess liquid is blotted off and the stain is allowed to dry. This procedure will destroy the original specimen (such as a protein complex), but leave an imprint of its shape. Negative stain limits resolution due to the fact that electrons are not scattered by the specimen itself but by the metal stain around it. This means resolution cannot progress past the grain size of the stain (e.g. 20 Å in case of Uranyl acetate), which is vastly inferior to resolutions obtainable by cryo electron microscopy, where the signal derives from the specimen itself. Destruction of the original specimen can also induce structural artefacts,

such as flattening and other kinds of deformations. Nevertheless the fast and easy preparation coupled with unparalleled contrast makes negative staining popular in the EM field. Stains are usually selected based on their grain size and pH. Popular stains include Uranyl acetate and -formate, ammonium molybdate, as well as vanadium and tungsten based stains.

2.6.3 Cryo-EM:

In cryo-EM the specimen are embedded in a thin layer of vitreous ice to protect them from beam damage. As the ice is made from the solution the specimen is dissolved in, such as a buffer in case of protein complexes, cryo-EM provides structural information in a near-native environment, forgoing possible artefacts caused by negative staining. The trade-off is that electrons have to be scattered by the specimen themselves, which in the case of protein complexes are of almost the same density as the ice surrounding them (i.e. 1.3 g/cm^3 for protein vs. 0.9 g/cm^3 for ice). This results in severely reduced contrast, which can make it difficult to tell small specimen from the background noise. It does however allow for higher resolution. During sample preparation, the grid carrying the specimen is usually plunged into liquid ethane, which can rapidly cool down samples to -160°C at $\sim 100000^\circ\text{C/s}$, preventing the formation of ice crystals that could damage the specimen. The vitreous ice should not be much thicker than the specimen itself, as a thick ice layer can impede the penetration by electrons. Shortcomings of cryo-EM compared to negative staining include the requirement for more concentrated protein samples, longer and more laborious preparation and handling of vitrified material, as well as a lesser degree of protection from beam damage. In general samples should only be illuminated once and with an electron dose equal to or lower than $20 \text{ e}^-/\text{\AA}^2$ to prevent ionization and movement of particles during the imaging process.

2.6.4 CTF correction and filtering:

For processing of electron microscopy data, rectangular images of single particles are extracted from micrographs in a manual or automated fashion. The contrast transfer function (CTF) partially distorts information included in these particles, so this has to be corrected for. In electron microscopy the image contrast $I(k)$ is a product of the object transform $O(k)$ and the contrast transfer function $H(k)$, e.g. $I(k) = O(k)H(k)$. According to the equation

$$H(k) = -\sin\left[\left(\frac{\pi}{2}\right) C_s \lambda^3 k^4 + \pi \Delta f \lambda k^2\right]$$

the CTF is influenced by parameters such as the electron wavelength λ as defined by the electron source, the spherical aberration coefficient C_s as defined by the microscope lenses, the spacial frequency k and the defocus Δf of the picture. The latter is the only changing parameter as all the others are microscope-specific. By determining the defocus of every micrograph and filling in the microscope-specific parameters, micrographs can be CTF-corrected, thereby enabling the extraction of undistorted information from particles. In the next step, these particles are usually masked to exclude much of the particle surrounding from the analysis, reducing the noise level. Particles are also band pass filtered between frequencies corresponding to the Nyquist limit (i.e. twice the sampling rate of the device used for recording micrographs) and the maximum diameter of the particle, thereby excluding frequencies that do not carry information corresponding to the particle itself. Filtering is performed in Fourier space, with low frequencies corresponding to coarse features and high frequencies corresponding to the finer details of particles. As processing methods often rely on cross correlation between images, all particles are normalized to the same grey values. This helps to avoid potential follies caused by large cross correlation peaks based solely on high numerical grey scale values.

2.6.5 Multivariate statistical analysis:

To find repeating motifs or differences within a dataset of single particles, they can be subjected to multivariate statistical analysis (MSA). This is also being referred to as reference-free classification, as it does not use outside information, just what can be obtained from the dataset itself. MSA classifies data based on eigenimage analysis. Eigenimages are created in a process that starts by concatenating the rows of pixels in each particle image, generating 1-dimensional vectors with as many elements as pixels in a single image. These vectors are combined into a matrix (M) and the mean is subtracted from each image. Eigenvectors are calculated from the covariance matrix (C), which is a product of the matrix M and its transposed counterpart M^T , as in $C = MM^T$. Since these eigenvectors contain as many pixels as the original images they are also referred to as eigenimages. An N -dimensional covariance matrix will result in N^2 eigenimages, so only the eigenimages with the highest associated eigenvalues are kept for further analysis. They are referred to as the principal components, as they represent the most common variations inside the dataset. In the images, strong variations will show up as white or black areas, indicating for example alternating subunits inside a symmetric arrangement, missing subunits or simply the difference between various views of the same object.

2.6.6 Angular reconstitution:

Unless a reference structure of the particle is available, initial 3D models have to be calculated *ab-initio*. Conical tilt experiments (Radermacher et al. 1987) can generate 3D models from pairs of tilted and untilted particles by making use of the known tilt angle and axis, enabling the correct assignment of Euler angles. The common lines method uses a different approach, exploiting the fact that all 2D projections of the same 3D object have at least one 1D line in common. This can be applied to both Fourier (Fuller 1987) and real space (Van Heel 1987). The latter works through the generation of sinograms, which are an assortment of 1D projections of a 2D image over a range of 360° . By mapping the cross correlation between two sinograms over the full angular range, the corresponding angle between the two original projections can be found. Adding more projections and finding their corresponding angles fills up the angular space of the model step by step, resulting in an initial 3D reconstruction of the desired object. The process is therefore also referred to as angular reconstitution. Both conical tilt and angular reconstitution are usually performed with a small number of particles exhibiting high quality and contrast.

2.6.7 Projection matching:

In order to improve an existing initial model a larger dataset of particles is needed. These particles are aligned to the reference model through a process called projection matching. Re-projections of the existing 3-dimensional structure are generated to fill the entire angular space of $\beta = 0-180^\circ$ and $\gamma = 0-360^\circ$ (or smaller angular spaces in case of symmetric particles) and each particle is compared to each re-projection by means of cross correlation. This will result in each particle being assigned the most probable shifts in x,y direction and rotational parameters for α , β and γ according to the reference projections. Projection matching is an iterative process and will be repeated with each round of 3-dimensional reconstructions, resulting in a refinement of particle shifts and angles until they are stable. Independent validation of the resulting model can encompass a variety of different techniques both on a 2D or 3D level, like fitting existing crystal structures into the density or comparing it to homologous models. In EM the most common way of validation is performing MSA on the images that were aligned to the 3D reference, checking if the resulting class averages (i.e. the projections) resemble the re-projections of the 3D structure.

3. Materials:

3.1 Chemicals:

Unless stated otherwise, all chemicals were purchased from Carl Roth, Sigma Aldrich, Merck, Serva, Fluka or VWR. Cell culture media components were purchased from Formedium.

3.2 Plasmids and strains:

3.2.1 Plasmids:

Name	Insert	derived from	Source
pJD657	$P_{Cup1-2myc-Pba1-T_{Pba1}}$ $P_{Pba2-Pba2-T_{Pba2}}$ $P_{Pba3-Pba3-T_{Pba3}}$ $P_{Pba4-Pba4-2xHA-T_{cyc1}}$	YCplac33 (<i>SacI</i> + <i>HindIII</i>)	J. Dohmen
pMO3	$5' \Delta-pre1-1-Flag-6His-T_{Pre1}$	YIplac211 (<i>HindIII</i> + <i>EcoRI</i>)/ <i>SacI</i>	M. Nunes
pMO4	$5' \Delta-pre1-Flag-6His-T_{Pre1}$	YIplac211 (<i>HindIII</i> + <i>EcoRI</i>)/ <i>SacI</i>	M. Nunes
YIp5 Ump1-GFP-HA-TEV-ProA	<i>Ump1-GFP-2xHA-TEV-2xProA-URA3</i>	YIp5 Ump1-GFP-HA	C. Enenkel

3.2.2 *E. coli* strains:

Strain	Genotype	Source
DH5 α	F ⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ 80dlacZ Δ M15 Δ (<i>lacZYA-argF</i>)U169, hsdR17(r _K ⁻ m _K ⁺), λ -	Sambrook <i>et al.</i> 1989

3.2.3 *S. cerevisiae* strains:

Strain	Genotype	derived from	Source
AM31	<i>MATa pre4ΔC19 F6H-UMP1 blm10Δ::kanMX6</i>	JD47-13C	Marques <i>et al.</i> 2007
MO23	<i>MATa pre1-1F6H::YIplac211</i>	JD47-13C	M. Nunes
MO24	<i>MATa PRE1-F6H::YIplac211</i>	JD47-13C	M. Nunes
MO27	<i>MATa P_{GALI} PBA1::TRP1 P_{GALI} PBA2::HIS3</i>	AM31	M. Nunes
BMF1	<i>MATa1 blm10Δ::HIS3</i>	WCGa	Fehlker <i>et al.</i> 2003

4. Methods:

4.1 Cell culture methods:

4.1.1 Media for *S. cerevisiae* cell cultures:

Complete yeast medium (YPD): 10 g yeast extract

20 g peptone

20 g D-glucose

add dH₂O to final volume of 1 L

Minimal medium (his⁻ ura⁻ leu⁻): 1.9 g yeast nitrogen base (without amino acids/(NH₄)₂SO₄)

5 g (NH₄)₂SO₄

20 g D-glucose

0.65 g dropout powder (his⁻ ura⁻ leu⁻)

3 ml 2 M NaOH

add dH₂O to final volume of 1 L

To abolish selection markers the following media

components were added:

30 mg/L histidine

30 mg/L uracil

100 mg/L leucine

Agar plates were prepared same as the medium but with an additional 20 g of agar per litre.

4.1.2 Medium for *E. coli* cell cultures:

Luria broth (LB): 5 g yeast extract

10 g tryptone

10 g NaCl

add dH₂O to final volume of 1 L

Agar plates were prepared same as the medium but with an additional 15 g of agar per litre.

Ampicillin was added to a final concentration of 100 µg/ml before pouring the plates.

4.1.3 *S. cerevisiae* cell cultures:

Pre-cultures of 50 ml minimal medium were inoculated with plate-grown yeast cells and shaken at 120 rpm. After overnight incubation at room temperature cells were either harvested (e.g. for cell lysis) or transferred into larger cultures of YPD medium. These larger cultures were grown at room temperature and shaken at 120 rpm until an OD₆₀₀ of 3.0 - 5.0, at which point they were harvested ($1 \text{ OD}_{600} \approx 3 \cdot 10^7$ cells/ml).

4.1.4 *E. coli* cell cultures:

Single colonies of plate grown *E. coli* cells were used to inoculate 5 ml of LB medium supplemented with 100 µg/ml ampicillin. Cells were grown overnight at 37°C and shaken at 120 rpm before harvesting.

4.2 Molecular biology methods:

4.2.1 Plasmid Isolation:

Plasmid DNA was isolated from DH5α cells using the Qiagen Spin Miniprep kit and instructions provided by the manufacturer.

4.2.2 Restriction digestion:

Plasmid DNA isolated from 3 ml of DH5α cells was cleaved using the following procedure.

40 µl plasmid DNA (≈ 4 µg)

30 µl 10x buffer 3 (NEB)

4 µl *Bst*XI (NEB)

3 µl 10 mg/ml BSA (NEB)

223 µl sterile dH₂O

The sample was incubated for 3 h or overnight at 37°C. Cleavage products were confirmed by agarose gel electrophoresis.

4.2.3 Transformation of *S. cerevisiae* cells by electroporation:

To generate cells competent for DNA uptake, yeast cultures were grown in 50 ml YPD medium until the early logarithmic phase ($\text{OD}_{600} = 0.8\text{-}1.0$) and spun down for 5 min at 3000

rpm (3463 x g) inside a sterile 50 ml tube. The supernatant was discarded and cells were washed with 10 ml sterile dH₂O. After a second centrifugation step cells were resuspended in 10 ml ice cold 1 M sorbitol and spun down at 4°C. The pellet was mixed with 1 ml ice cold 1 M sorbitol and transferred to a sterile 1.5 ml tube. After a brief 60 s spin at 6000 rpm (8452 x g) and 4°C the supernatant was discarded. The pellet was resuspended in equal volumes of ice cold 1 M sorbitol and kept on ice until use. 40 µl of cell suspension were transferred to a BioRad Gene Pulser cuvette and briefly mixed with 4 µl of digested plasmid DNA (~400 ng). The sample was incubated on ice for 5 min. Electroporation was performed on a BioRad Gene Pulser using settings for 1500 V, 25 µF and 200 Ω. Immediately after the shock, 1 ml of ice cold 1 M sorbitol was added to the cells. The sample was transferred to a sterile 1.5 ml tube and spun down for 2 min at 6000 rpm (8452 x g) and 4°C. The supernatant was discarded. The pellet was resuspended in 80 µl of ice cold 1 M sorbitol and plated on agar plates with the appropriate selection marker.

This thesis used the general purpose *E. coli* vector YIp5 (Struhl 1979) to integrate DNA into the yeast genome through homologous recombination as described previously (Enenkel et al. 1998). After electroporation, His⁺ and Ura⁺ markers selected for the recombined locus once cells were plated on the respective selection plates.

4.2.4 Agarose gel electrophoresis:

Agarose gel electrophoresis can be used to separate nucleic acid strands of varying sizes inside an agarose matrix. Separation was performed using a BioRad Mini-Sub horizontal electrophoresis device and 1x TBE running buffer. For visualization of bands under UV light, DNA Stain G (Serva) was added in a 1:70000 dilution while casting the gels. Nucleic acid samples were mixed with 5x sample buffer before the run. Gels were run at 100 V for 1 h before documentation on a SafeLab Imager (Intas). Gene Ruler DNA ladder (Fermentas) was used to estimate sizes of nucleic acid chains.

10x TBE running buffer: 900 mM Tris-HCl pH 8.0

900 mM boric acid

10 mM EDTA

5x sample buffer: 10 % (v/v) 10x TBE running buffer

50 % (v/v) glycerol

Bromphenol blue

4.3 Protein biochemistry methods:

4.3.1 Protein precipitation using TCA and NaDOC:

Protein containing solutions up to 1 ml were mixed with 100 μ l 72 % (v/v) TCA and 100 μ l 0.15 % (w/v) NaDOC. After incubating for 20 min at room temperature (or overnight at 4°C), they were spun down at 13000 rpm (16060 x g) and 4°C for 10 min. The supernatant was discarded and pellets were washed with 0.5 ml ice cold acetone. Samples were again spun down for 5 min at 13000 rpm (16060 x g) and 4°C. The supernatant was discarded and remaining acetone was allowed to evaporate for 10 min under the fume hood. For analysis by SDS-PAGE the protein pellets were mixed with appropriate amounts of 1x Laemmli sample buffer.

4.3.2 SDS polyacrylamide gel electrophoresis (PAGE):

SDS-PAGE is used to separate proteins under denaturing conditions inside a polyacrylamide matrix based on their electrophoretic mobility. Gels were cast as described in table 2 following a modified protocol introduced by Laemmli (1970) and allowed to polymerize for 25 min. Before the run, pellets were resuspended in 1x Laemmli sample buffer and boiled at 95°C for 5 min. Separation was performed in a BioRad Mini-Protean Tetra Cell system using 1x Laemmli running buffer and 200 V constant voltage for 50 min. The molecular weight of unknown protein bands was estimated using broad range pre-stained protein marker (NEB).

Table 2: Composition of SDS-PAGE gels

Gel component	Stacking gel (5 %)	Separating gel (12 %)
dH ₂ O	3.4 ml	3.3 ml
Acrylamide solution (30 %) containing 37.5 : 1 acrylamide : bisacrylamide	0.83 ml	4 ml
1.5 M Tris-HCl pH 8.8	-	2.5 ml
1 M Tris-HCl pH 6.8	0.64 ml	-
10 % (w/v) SDS	50 μ l	100 μ l
10 % (w/v) Ammonium persulfate	50 μ l	100 μ l
TEMED	5 μ l	10 μ l

2x Laemmli sample buffer: 126 mM Tris-HCl pH 6.8
20 % (v/v) glycerol
4 % (w/v) SDS
5 % (v/v) β -mercaptoethanol
Bromphenol blue

10x Laemmli running buffer: 250 mM Tris
2 M glycine
1 % (w/v) SDS

4.3.3 Coomassie Brilliant Blue stain:

Staining of SDS-PAGE gels with Coomassie Brilliant Blue represents the most common way of visualizing proteins after electrophoretic separation. The stain can bind basic amino acid side chains and is therefore suitable for unspecific staining of all proteins. Gels were shaken in a dish containing Coomassie staining solution for 60 min at room temperature before the gel was transferred to destain. Destaining was continued until excess Coomassie Brilliant Blue was removed from the gel and bands were clearly visible.

Coomassie staining solution: 0.25 % (w/v) Coomassie Brilliant Blue R-250
45 % (v/v) MeOH
10 % acetic acid

Destaining solution: 45 % (v/v) MeOH
10 % acetic acid

4.3.4 Amido black stain:

Amido black staining was performed on Western blot membranes after protein transfer onto the membrane. The dye stains all proteins unspecifically, enabling a visual control of protein transfer efficiency. The membrane was incubated in 45 ml dH₂O mixed with 5 ml of amido black staining solution for 5 min at room temperature. Excess amido black was washed off with dH₂O and an image of the membrane was taken.

Amido black staining solution: 0.1 % (w/v) amido black
40 % MeOH
10 % acetic acid

4.3.5 Silver stain:

Silver stain is used to visualize protein bands in SDS gels after electrophoretic separation (e.g. SDS-PAGE). In contrast to Coomassie-based stains, silver stain has a much higher sensitivity, being able to visualize up to 1 ng of protein (compared to a lower limit of 50 ng in Coomassie Brilliant Blue G-250).

To silver stain proteins, a polyacrylamide gel was treated as follows:

- 1) Washed with 50 % (v/v) EtOH for 20 min
- 2) Washed with 5 % (v/v) EtOH for 20 min
- 3) Reduced proteins with 35 μ M DTT in dH₂O for 5 min
- 4) Stained proteins with silver nitrate solution for 10 min
- 5) Rinsed twice with dH₂O
- 6) Rinsed twice with developing solution
- 7) Added more developing solution and developed bands until the desired darkness
- 8) Stopped the reaction by adding several spoonful of citric acid monohydrate until the fizzing subsided
- 9) Incubated for another 5 min, then transferred to dH₂O

Silver nitrate solution: 0.1 % (w/v) silver nitrate
0.0037 % formaldehyde (added fresh)

Developing solution: 3 % (w/v) sodium carbonate
0.0185 % formaldehyde (added fresh)
chilled to 4°C

4.3.6 Western blot:

Western blots are used for immunological detection of specific proteins after they are separated by electrophoresis. In this thesis, Western blotting was performed in a semi-dry way, meaning that the membrane and blotting papers were only briefly soaked in transfer

buffer before use. Gels were transferred to the blotting station (PeqLab) and sandwiched between a nitrocellulose membrane (Whatman) and 2 layers of blotting paper. A maximum current of 400 mA was applied for 1.5 h before the membrane was transferred to an amido black solution for staining (see method 4.3.4). To visualize fusion proteins coupled to protein A moieties by chemiluminescence, the membrane was treated as follows:

- 1) Incubation in blocking solution for 25 min at room temperature
- 2) Staining with 1:10000 dilution of α -mouse HRP-conjugated antibody (from rabbit; Santa Cruz) in blocking solution for 2 h at 4°C
- 3) Washing with TST for 5 min at room temperature (repeat 2x)
- 4) Incubating with ECL developing solution for 1 min, after which chemiluminescence was captured on film (Hyperfilm ECL, Amersham)

Transfer buffer: 12.5 mM Tris
100 mM glycine
20 % (v/v) MeOH
0.05 % (w/v) SDS

Blocking solution: 5 % (w/v) milk powder in dH₂O

TST: 10 mM Tris-HCl pH 8.0
150 mM NaCl
0.01 % (v/v) Tween 20

ECL developing solution: 100 mM Tris-HCl pH 8.5
0.198 mM coumaric acid
1.25 mM luminol
0.01 % (v/v) H₂O₂

4.3.7 Cell lysis by bead beater:

The bead beater was used to lyse cells for purification of native protein complexes. The cells pellet was thawed in equal amounts of lysis buffer and transferred to a 50 ml tube. For optimal results, glass beads (0.25-0.5 mm diameter, Carl Roth GmbH) were added equalling 1.5 x the volume of the cell suspension. The cells were lysed using a Fastprep24 bench top homogenizer (MP Biomedicals) at 6.5 m/s for 30 s per cycle. The number of cycles depended

on the size of the sample. In general >90 % of cells were lysed after 8-10 cycles. Cell lysis was monitored by checking the sample under a light microscope. Lysate was separated from glass beads by poking holes into the bottom of the tube, attaching a second tube and centrifuging for 4 min at 1000 rpm (219 x g).

4.3.8 Non-native cell lysis:

Non-native lysis was performed using the procedure introduced by Yaffe and Schatz (Yaffe & Schatz 1984) for quick protein extraction and Western blot analysis. Cells from 3-5 ml of culture were harvested by centrifugation at 13000 rpm (16060 x g) for 2 min. The pellet was resuspended in 1 ml of dH₂O and mixed with 160 µl 1.85 M NaOH as well as 80 µl β-mercaptoethanol. After incubation for 10 min on ice, 160 µl of 50 % (v/v) TCA were added and the sample was again incubated on ice for 10 min. The precipitate was spun down at 13000 rpm (18312 x g) for 10 min, the supernatant was discarded and the pellet was washed with 0.5 ml ice cold acetone. After another centrifugation step the supernatant was again discarded and the remaining acetone was allowed to evaporate for 10 min under the fume hood. For analysis by SDS-PAGE and Western blot the pellet was resuspended in 150 µl 1x Laemmli sample buffer.

4.3.9 Purification of 15S^{GFP} complexes for electron microscopy:

15S^{GFP} proteasome precursor complexes were affinity purified via the protein A tag C-terminally attached to Ump1-GFP using an IgG sepharose matrix (GE Life Science). Cells were lysed as described in method 4.3.7 using buffer T (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 % (v/v) glycerol, 0.5 mM DTT) supplemented with 0.002 % (w/v) DNase I (Applichem) and 1 protease inhibitor tablet (Roche) as lysis buffer. The lysate was spun down at 14000 rpm (23426 x g) and 4°C for 25 min. The supernatant was applied to IgG sepharose beads equilibrated with 2x 5 column volumes TST, 2x 5 column volumes 0.5 M HAc pH 3.4 and 5x 10 column volumes buffer T. After incubation for 2 h at 4°C, the flow through was collected and the column was washed with 3x 20 column volumes of buffer T. The IgG sepharose beads were incubated with 2 column volumes of buffer T and ~37 µg TEV protease for 2 h at room temperature to elute 15S complexes from the column material. To further purify the 15S complex and separate TEV protease from the protein mixture the eluate was applied to a 10-40 % glycerol gradient as described in method 4.3.11. All fractions of the

gradient were investigated by SDS-PAGE and 15S^{GFP} containing samples were used for negative stain electron microscopy.

Identification of precursor-specific proteins inside Coomassie stained bands of SDS-gels was carried out by Dr. Thomas Fröhlich in the laboratory for functional genome analysis (LAFUGA, Gene Center Munich).

4.3.10 Purification of 15S^{GFP} complexes for DSS cross-linking experiments:

As DSS cross-linking of proteins is incompatible with Tris buffer due to its primary amino group, the 15S^{GFP} proteasome precursor was also purified in buffer H (25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 5 % (v/v) glycerol, 0.5 mM DTT) containing HEPES as the buffering component. The procedure is in large parts analogous to the one described above with buffer H substituting for buffer T. After TEV cleavage the sample was cross-linked with DSS H12/D12 (see method 4.3.13) and further purified by a 10-40 % glycerol density gradient (see method 4.3.11). The fractions containing cross-linked 15S^{GFP} precursor complexes were pooled and concentrated using Amicon Ultra centrifugal filter units with a 100 kDa cutoff (Millipore). The concentrated sample was used for analysis by mass spectrometry.

4.3.11 Analytical ultracentrifugation:

Analytical ultracentrifugation refers to a method of separating proteins or protein complexes of different mass and shape by use of centrifugal force. During centrifugation proteins are forced through a gradient of increasing density, which creates drag and separates proteins according to their sedimentation behaviour. In this thesis density gradients comprising of 10-40 % glycerol were used. Stock solutions containing the respective buffer supplemented with 10 % or 40 % glycerol were prepared prior to centrifugation. 6 ml of each buffer solution were mixed using a Biocomp gradient station. A maximum of 600 µl sample were applied to the top of the gradient and tubes were spun for 16 h at 40000 rpm (283807 x g) and 4°C using a SW40 Ti rotor and a Optima L-80 ultracentrifuge (Beckman). For further analysis the gradient was manually separated into 600 µl fractions starting from the top (light fractions) to the bottom (heavy fractions).

4.3.12 Determination of protein concentrations:

Protein concentrations were determined using the bicinchoninic acid assay (BCA, Thermo Scientific), which relies on colour changes from green to purple to signify increasing protein content of a sample solution. Peptide bonds of the amino acid chain reduce Cu^{2+} ions to Cu^+ , which are in turn chelated by bicinchoninic acid, forming a purple coloured complex.

The working reagent was prepared as per the manufacturer's instructions and 200 μl were mixed in a microplate well with 10 μl of the sample, followed by incubation for 30 min at 37°C. The absorbance at 562 nm was measured on a Tecan Infinite M1000 microplate reader. Samples of unknown concentration were compared to standards consisting of bovine serum albumin (Sigma Aldrich) solutions ranging from 20 $\mu\text{g/ml}$ to 2000 $\mu\text{g/ml}$. Concentrations were deduced using linear regression.

4.3.13 Protein cross-linking using disuccinimidyl-suberate (DSS):

To elucidate the subunit composition of protein complexes, individual components can be chemically cross-linked to one another. Coupled with a proteolytic digestion and mass spectrometric analysis of the resulting peptides, spacial restraints can be derived that allow insights into neighbouring subunits as well as their orientation inside the complex. DSS can be used to cross-link primary amino groups found in lysine side chains and the protein N-terminus (see figure 6). The arm length of DSS is 11.4 Å. Together with the lysine side chains this results in a maximum distance of ~30 Å between lysine $\text{C}\alpha$ atoms.

The amount of protein used for cross-linking varied depending on the purpose of the reaction. To titrate the ideal protein-to-cross-linker-ratio, 1.5-2.0 μg protein samples were cross-linked and then investigated by silver stain. For the final mass spectrometric analysis, 40-50 μg of protein were used.

Instead of generic DSS, for this method a mixture of DSS isotopes ($\text{H}12/\text{D}12$; Creative Molecules Inc.) is used, in which 50 % of all molecules contain deuterated side chains (see figure 6). To calculate the amount of DSS $\text{H}12/\text{D}12$ necessary for efficient cross-linking, the lysine content of the sample was determined. As a rule of thumb, 1 μg of protein contains ~500 pmol of lysines. To reconstitute the DSS $\text{H}12/\text{D}12$, 1 mg was mixed with 53 μl of N,N-dimethylformamide (DMF), resulting in a 25 mM stock solution. Protein samples were incubated with the appropriate amount of reconstituted DSS $\text{H}12/\text{D}12$ for 30 min at 30°C. The cross-linking reaction was quenched by adding ammonium bicarbonate to a final

concentration of 100 mM. Samples were again incubated for 30 min at 30°C before storage at 4°C.

Protein cross-linking was used on the 15S and 15S^{GFP} proteasome precursor complexes. Mass spectrometric analysis of cross-linked samples was carried out by the lab of Dr. Franz Herzog (Gene Center Munich) as described previously (Leitner et al. 2012; Herzog et al. 2012).

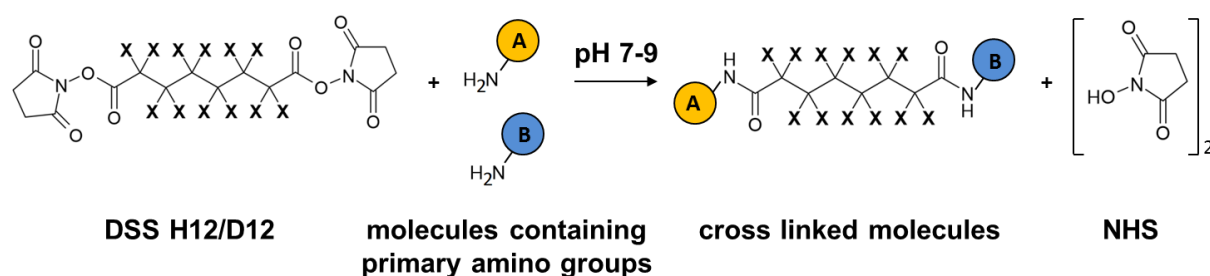


Figure 6: Cross-linking reaction between DSS H12/D12 and molecules containing primary amino groups. During the reaction, the N-hydroxysuccinimide (NHS) groups are split from the DSS H12/D12 and amide bonds can be formed between the cross-linker and primary amino groups of other molecules. The X in DSS H12/D12 denotes the possibility of this molecule to either contain hydrogen or deuterium in these places.

4.4 Electron microscopy methods:

4.4.1 Glow discharging of EM grids:

Glow discharging deposits charged particles on the surface of otherwise hydrophobic electron microscopy grids, making it easier for proteins to bind to this surface. Grids were placed inside the plasma cleaner (Harrick Plasma) on a glass dish and the vacuum chamber was allowed to settle at a pressure level of $2.2 \cdot 10^{-1}$ Torr using atmospheric gas. Grids were glow discharged for 30 s, then stored in a sealed container and used within a timeframe of 60 min.

4.4.2 Negative stain:

All negative stain experiments were carried out using Uranyl acetate (Ted Pella) as heavy metal stain. 3.5 µl of sample containing ~50 µg/ml protein were applied to a glow discharged grid with continuous carbon surface (Quantifoil or Plano Cu 400 mesh) and incubated for 45 s. Excess liquid was blotted off using Whatman #1 blotting paper. 3.5 µl of 2 % (w/v) Uranyl acetate solution were applied to the grid and incubated for 15 s. The surface of the grid was then incubated in four 25 µl drops of 2 % (w/v) Uranyl acetate for 10 s each. After the last drop, excess liquid was again blotted off and the grid was allowed to dry for 5 min. Grids were stored at room temperature in the dark.

4.4.3 Preparation of vitrified samples for cryo-EM:

Grids with vitrified samples for cryo-EM were prepared using a Mark IV Vitrobot (FEI) with a temperature and humidity controlled chamber. Before the start of the procedure the chamber was allowed to settle at 4°C and 100 % humidity. Grids (Quantifoil R 2/2 + 2 nm carbon on a Cu 300 mesh) were treated with 5 drops of chloroform and allowed to dry completely before glow discharging to dispose of residual plastic coating material. 3.5 µl of sample with a protein concentration of 125 µg/ml were applied to a glow discharged grid and incubated in the chamber for 45 s. Excess liquid was blotted off using 2 layers of Whatman #1 paper for 2-3 s and a blot force of 0. The grid was then plunged into liquid ethane without additional drain time. Grids were stored in liquid nitrogen until use.

4.4.4 Nanogold labelling:

Nanogold can be used for the labelling of individual subunits of a protein complex. Due to the high electron density of the gold particle, it is easily visible in both negative stain and cryo-EM images as black dots. In this thesis 5 nm Nanogold particles (Nanoprobes) coated with Ni²⁺ NTA (nitrilotriacetic acid) were used to bind proteins carrying a his₆ tag. Solutions containing 50 µg/ml of the desired protein complex were mixed with 10 nmol/ml Nanogold particles at a ratio of 50:1 and incubated on ice for 10 min. Proteins were then stained with Uranyl acetate as described in method 4.4.2.

4.4.5 Collection of negative stain data:

For the purpose of screening negative stain grids, a FEI Morgagni transmission electron microscope (BioCenter Munich EM facility, Martinsried) equipped with an SIS Megaview 1K CCD camera and running at 80 kV was used. Images were collected at a nominal magnification of 60000x. Micrographs of negatively stained complexes for 3D reconstructions were collected on a FEI Tecnai G2 Spirit (BioCenter Munich EM facility, Martinsried) using a magnification of 96000x at specimen level and an Eagle CCD camera with 2048 x 2048 pixels. The microscope utilized a lanthanum hexaboride (LaB₆) electron source set to an extraction voltage of 120 kV. For imaging, the electron dose was set to 20 e⁻/Å². Images were collected at defoci between -0.3 µm and -1.0 µm.

4.4.6 Collection of cryo-EM data:

Micrographs of cryogenically frozen particles were collected on a FEI Tecnai G2 Spirit microscope (see above for specifications) or a FEI Tecnai F20 microscope (Max Planck Institute for Biochemistry EM facility, Martinsried) using a magnification of 84270x at specimen level and a CCD camera with 4096 x 4096 pixels. The Tecnai F20 microscope utilized a field emission gun electron source set to an extraction voltage of 200 kV. The images were taken under low dose conditions at an electron dose of $20 \text{ e}^-/\text{\AA}^2$. Cryo-EM images on both microscopes were collected at a defocus range between -1.5 μm and -4.0 μm .

4.4.7 Pre-processing of electron microscopy images:

The contrast transfer function (CTF) of micrographs was determined using CTFFIND3 (Mindell & Grigorieff 2003), and micrographs were CTF corrected using SPIDER (Frank et al. 1996). Negatively stained 15S particles were picked manually using the program BOXER, which is part of the EMAN1 suite (Ludtke et al. 1999). Particles from cryo-EM images and negatively stained 20S *pre1-1* were picked in an automated fashion using Find-EM (Roseman 2004) and class averages of handpicked data as reference images. Picked particles were written into stacks using EMAN1 BatchBoxer or LABEL (Crowther et al. 1996). Cryo-EM data of 15S complexes as well as negative stain data of 20S *pre1-1* complexes was sorted using the z-score function of XMIPP (Sorzano et al. 2004) and bad particles were excluded based on visual inspection. Remaining particles were masked, band-pass filtered and normalized in IMAGIC (van Heel et al. 1996) using the command *incore-prepare-filtered* and centred in an iterative fashion using *summer.e*, *rotatrim.e*, *alidir.e* and *alium.e*.

4.4.8 Multivariate statistical analysis (MSA) and multi reference alignments (MRA):

Particle stacks were subjected to multiple consecutive rounds of MSA and MRA in IMAGIC to further assess the quality of data. MSA was performed using *msa-run*, *msa-class* and *msa-sum*. Data sets were classified grouping ~10 particles/class for negative stain data and ~20 particles/class for cryo-EM data. Visual inspection of class averages lead to exclusion of bad particles on a per-class basis (class sorting). In the case of the 20S *pre1-1* complex, class sorting was also used to further sort the particles into 20S *pre1-1* bound to one, two or no Pba1-Pba2 heterodimers. For the MRA, references were generated from class averages of the dataset using the commands *align-mass-centre* for mass centring, *arithmetic-with-images* for masking and *norm-var* for normalization. The MRA was run using *multi-reference-alignment*

with aforementioned references, permitting full 360° rotations and shifts equal to 20 % of the box size.

4.4.9 Angular reconstitution:

Angular reconstitution in IMAGIC was used to create initial models of 15S, 20S *pre1-1* and 20S utilizing negative stain class averages of top- and side-views gained during the aforementioned combined MSA and MRA procedure. Before the start of the procedure these class averages were visually inspected for mismatched particles. Angular reconstitution was executed using the command *euler*. Since the 15S complex has a pseudo 7-fold symmetry, C7 symmetry was enforced during angular reconstitution to increase the reliability of the initial 3D reconstruction. In the case of 20S *pre1-1* no symmetry was applied, as the initial 3D reconstruction was made of 20S *pre1-1* bound to only a single Pba1-Pba2 heterodimer. For the mature 20S, C2 symmetry was applied to the initial 3D reconstruction. To check the quality of the angular reconstitution, projections were compared to re-projections of the initial model generated with the command *true-threed*. If they were largely a match the model was considered reliable.

4.4.10 Refinement:

Refinement of the 3D model was mainly performed in SPIDER. Initially the angular spacing for re-projections was set at 8° and gradually reduced to 4° (or 2° in the case of 20S *pre1-1*) over the course of the refinement process. Re-projections were generated over a range of $\beta = 0-180^\circ$ and $\gamma = 0-360^\circ$ (or $\gamma = 0-180^\circ$ in the case of mature 20S) using the command *PJ 3Q*. The images were aligned to the references using the command *AP SH* with the mirror option disabled. Three-dimensional models were reconstructed in 80 iterations using *BP RP*. Even though 20S *pre1-1* is in principal a particle exhibiting C2 symmetry, no symmetry was assumed because most of the 20S *pre1-1* particles only had one Pba1-Pba2 heterodimer bound. For the mature 20S, C2 symmetry was assumed for the *BP RP* procedure. Only the top 75 % of particles with the highest alignment cross correlation score were used in the reconstruction. Each reconstruction was multiplied with a binary mask in IMAGIC using *threed-two-volumes-operations* to reduce the noise level. The refinement was continued until angles were considered stable (i.e. ≥ 90 % of all angles did not change compared to the previous refinement round). To assess the quality of the refinement and the resulting model, aligned particle images were subjected to independent classification by MSA and the resulting

class averages were compared to re-projections of the model. If the class averages (i.e. the projections) matched the re-projections, the refinement process was considered to be reliable. 3D reconstructions were visualized using CHIMERA (Pettersen et al. 2004) and PYMOL (DeLano 2002).

4.4.11 Resolution determination:

To determine the resolution of a reconstruction the data was split into odd and evenly numbered particles and reconstructions were performed using this data and the SPIDER command *BP RP*. The Fourier shell correlation (FSC) between the two resulting volumes was calculated using *RF 3*. The FSC was plotted against the normalized frequency and the resolution was read out at $\text{FSC} = 0.5$.

4.4.12 Determination of handedness:

As all EM reconstructions were performed using initial references of unknown handedness, the latter had to be determined independently. The most common method is using pairs of tilted and untilted particles (Rosenthal & Henderson 2003; Henderson et al. 2011). While the method was developed for cryo-EM data, it can also be applied to negatively stained particles. The same particles are recorded twice, once with an untilted grid and once with the goniometer tilted at a fixed angle. Since the difference in tilt angles is known to the experimenter, the corresponding difference in the particles Euler angles ϕ , θ and ψ can be calculated. The particles are aligned to the reference EM model of unknown handedness and the alignment output (i.e. the particles Euler angles) is compared to the expected outcome. If tilt pairs were aligned to the model with the right handedness, the alignment output would cluster around the goniometer tilt axis and the known tilt angle. Consequentially, using the wrong handedness would lead to a significant deviation of calculated particle Euler angles and the expected outcome. To generate reliable data, only particles that are tilted “in-plane” should be taken into account. Particles are considered to be tilted in-plane when their calculated tilt axis is an approximation of the real goniometer tilt axis. Particles are considered “out-of-plane” when one or both members of the tilt pair deviates too much from said axis. The latter also provides a measure of confidence for the alignment, as a high number of out-of-plane particles indicates difficulties in aligning particles to the reference model.

To generate tilt pairs, negatively stained particles were recorded at tilt angles of 0° and 15° . The CTF and the real tilt angle were calculated using CTFTILT (Mindell & Grigorieff 2003).

Corresponding untilted and tilted particles were picked using WEB. The contrast on particle images was inverted and they were aligned to the EM reconstruction of unknown handedness using FREALIGN (Grigorieff 2007). The angular difference between tilted and untilted particles was compared using TILTDIFF (Henderson et al. 2011) and plotted in a separate file.

Alternative methods for determining the handedness of a 3D reconstruction include using special restraints generated by cross-linking adjacent lysine side chains (as described in section 4.3.13) for fitting of available crystal structures. This method may exclude one of the possible hands in a process of elimination, for example if only one of the hands can accommodate all cross-links while the other cannot.

5. Biochemical results:

5.1 Transformation and expression tests:

Strains and plasmids necessary for the purification of 15S^{GFP} complexes were acquired from Prof. Dr. Cordula Enenkel (University of Toronto). *S. cerevisiae* was successfully transformed with linearized plasmid DNA encoding the construct *Ump1-GFP-2xHA-TEV-2xProA-URA*. All clones tested positive for expression of a fusion protein of the size of *Ump1-GFP-2xHA-TEV-2xProA* in α -ProA Western blots (see figure 7).

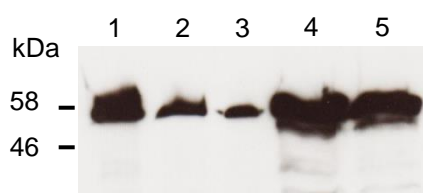


Figure 7: Total protein extracts of the strain BMF1, which was transformed with *BstXI*-digested *Ylp5 Ump1-GFP-2xHA-TEV-2xProA-URA* and tested against α -ProA in a Western blot analysis. Clones are indicated by numbers above the blot. The expected size of the fusion protein is 58 kDa.

5.2 Purification of proteasome precursor complexes:

All 15S precursor complexes used in this thesis were purified from strains carrying a Blm10 deletion to exploit the observation that deleting this non-essential activator can increase the homogeneity of protein preparations by excluding precursor complexes bound to Blm10 instead of Pba1-Pba2. Native 15S precursors (without the *Ump1-GFP-HA* fusion protein found in 15S^{GFP}) were purified from a strain carrying a C-terminal deletion of 19 amino acids in β 7, as this deletion also slows down dimerization of half-proteasomes and enriches 15S precursor complexes (Marques et al. 2007). In addition, this strain overexpressed Pba1-Pba2 in order to increase the occupancy of these chaperones. The 20S *pre1-1* complex was also purified from a strain overexpressing these chaperones.

Native 15S was affinity purified via a FLAG-his₆ tag on the N-terminus of *Ump1*, while 15S^{GFP} was affinity purified using a ProA tag on the C-terminus of the *Ump1-GFP-HA* fusion protein. 20S *pre1-1* and wild type 20S were purified via the FLAG-his₆ tag on the C-terminus of β 4. Table 3 gives an overview of all the proteasome precursor complexes used in this thesis and who purified them.

Table 3: Protein complexes investigated in this thesis and who purified them.

Protein complex	Purified by
15S	M. Nunes (Dohmen lab)
15S ^{GFP}	M. Kock (Wendler lab)
20S <i>pre1-1</i>	M. Nunes (Dohmen lab)
20S	M. Nunes (Dohmen lab)

Figure 8 shows the SDS-PAGE analysis of affinity purified 15S^{GFP} complexes. Individual subunits have been identified by peptide mass fingerprinting. All expected proteasome subunits are present in the complex, except for $\beta 5$ and $\beta 6$. These subunits were nevertheless identified during MS analysis of cross-linked 15S^{GFP} complexes (see section 5.5). The band pattern largely resembles the one previously found for proteasome precursor complexes affinity purified via Ump1 (see figure 1B in Marques et al. 2007 and figure 4B in Lehmann et al. 2010).

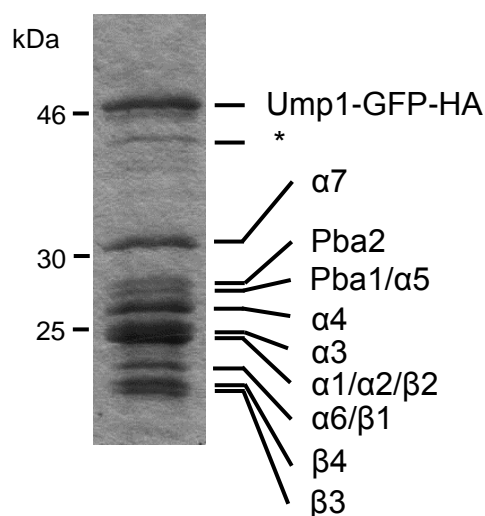


Figure 8: SDS-PAGE analysis and peptide mass fingerprinting of the affinity purified 15S^{GFP} complex. Individual subunits of the proteasome precursor complex are indicated on the right. The band marked with a star * was found to consist of a degradation product of the Ump1-GFP-HA fusion protein shown at 47 kDa.

The band at 47 kDa corresponds to the expected size of the Ump1-GFP-HA fusion protein after cleaving off the ProA moiety with TEV protease. A faint band (indicated by a star *) is visible slightly below the Ump1-GFP-HA band. This smaller band has been confirmed to also contain Ump1, hinting at partial degradation of the fusion protein on a low level. It is assumed that this degradation takes place on the N-terminus of Ump1, as the other end is fused to the more stable GFP. In order to check the molecular weight and structural integrity of the 15S^{GFP} complex it was subjected to analytical ultracentrifugation. Figure 9 shows gradient fractions of 15S^{GFP}. Fractions 1-3 are not included as they were not expected to contain any useful information. TEV protease and various low molecular weight contaminations are present in

fractions 4-5 at ~25 kDa, while the 15S^{GFP} complex is clearly visible in fractions 8-11. These fractions correspond to a molecular weight of ~400-500 kDa, which fits the molecular weight of 460 kDa calculated for the 15S^{GFP} complex. The band corresponding to partially degraded Ump1 already observed in figure 8 is again present in these fractions, indicating that it is still incorporated into the complex. Fractions 9-11 contain two bands (indicated by two stars ** and three stars ***) that have been confirmed by mass spectrometric analysis to correspond to Pba1 and Pba2. The two proteins are not present in fraction 8, reflecting the fact that they are not essential for assembly of the 15S complex and therefore not present in stoichiometric amounts. EM analysis confirms that fractions 8-11 mostly contain monomeric 15S^{GFP} complexes (see figure 12A).

Fractions 13-14 also contain 15S^{GFP} complexes. Their size corresponds to 700-900 kDa so it is likely that some 15S^{GFP} complexes form multimers (see also section 5.5). Fraction 21 corresponds to the high density end of the gradient, so all aggregates will be found here. The pattern of bands found in fraction 21 resembles that of the 15S^{GFP} protein pattern, hinting at the fact that a small number of complexes are unstable and prone to aggregation.

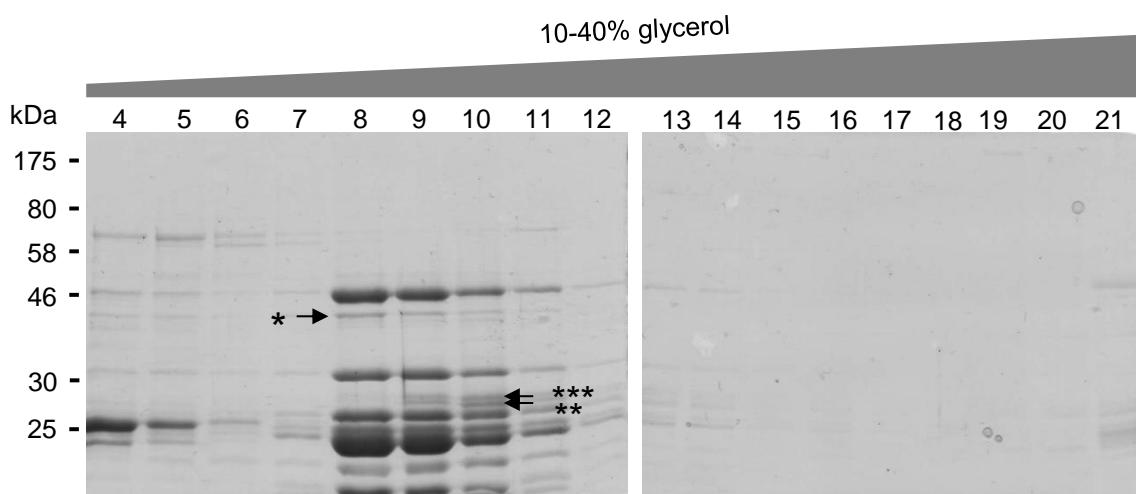


Figure 9: SDS-PAGE analysis of fractions derived from analytical ultracentrifugation of affinity purified 15S^{GFP} complexes on a 10-40% glycerol gradient. Gradient fractions (600 µl) are indicated by numbers. Fractions 1-3 were not included in the SDS-PAGE analysis. Partially degraded Ump1 is marked by a star * while the Pba1 and Pba2 are marked by two stars ** and three stars *** respectively.

5.3 Titrations for DSS H12/D12 cross-linking experiments:

In order to find the optimal concentrations for cross-linking 15S and 15S^{GFP} complexes with DSS H12/D12, they were incubated with increasing concentrations of the compound. Corresponding silver stained SDS-PAGE gels are shown in figure 10. The gel analyzing cross-linked 15S complexes shows a shift to highly cross-linked protein chains as DSS

H12/D12 concentration increases. This can be seen by increasing amounts of protein in higher molecular weight species near the top of the gel, while the lower molecular weight bands in the range of ~17-35 kDa vanish. For cross-linked 15S^{GFP} this shift is not as pronounced (possible due to small differences in protein concentration between the samples), but the overall pattern is similar to the 15S sample. Both native 15S and 15S^{GFP} complexes cross-link efficiently at 0.75-0.8 mM DSS H12/D12, so these concentrations were chosen for all subsequent cross-linking experiments.

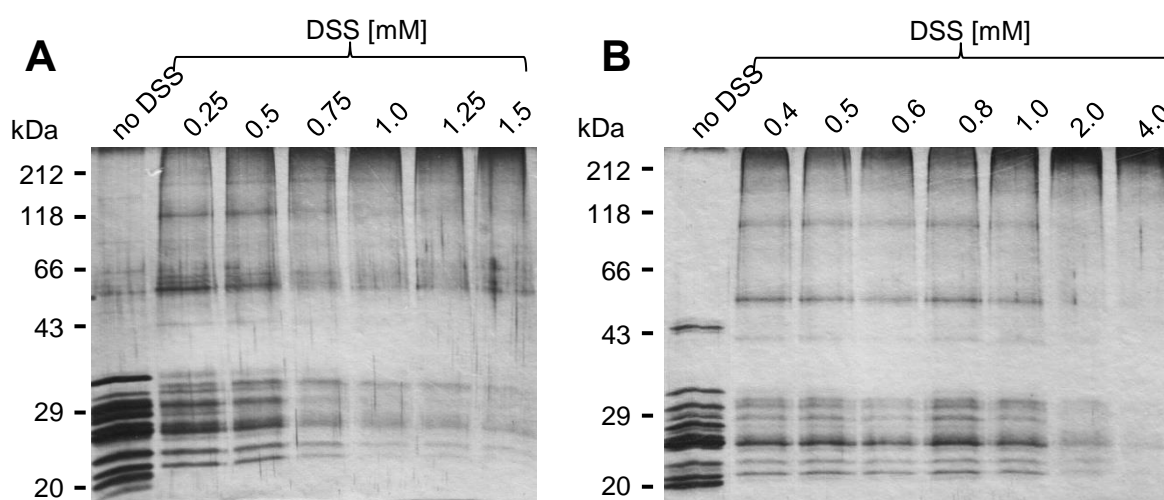


Figure 10: Silver stained SDS-PAGE gels of 15S complexes (A) and 15S^{GFP} complexes (B) titrated with increasing concentrations of DSS H12/D12. Each lane contains 2 µg of protein complex. The leftmost lanes of both gels show control samples not cross-linked with DSS H12/D12.

5.4 Analytical ultracentrifugation of cross-linked proteasome precursor complexes:

In order to enrich cross-linked monomeric complexes in the sample and exclude aggregates, cross-linked complexes were further purified using glycerol gradient ultracentrifugation. Analogues to the fractions shown in figure 9, figure 11 shows cross-linked 15S^{GFP} complexes after analytical ultracentrifugation. After incubation with DSS H12/D12, subunits have been cross-linked into high molecular weight amino acid chains, which mainly populate the upper parts of the SDS-gel. Fractions 8-11 still contain the bulk of protein, just like in the non-cross-linked sample. Only a small number of subunits in these fractions remain untouched by the cross-linking reaction, as evidenced by faint bands in the range of ~20-30 kDa. EM analysis confirms that these fractions correspond to cross-linked monomeric 15S^{GFP} complexes (figure 12A), hence they were used for further analysis using mass spectrometry (MS). The high

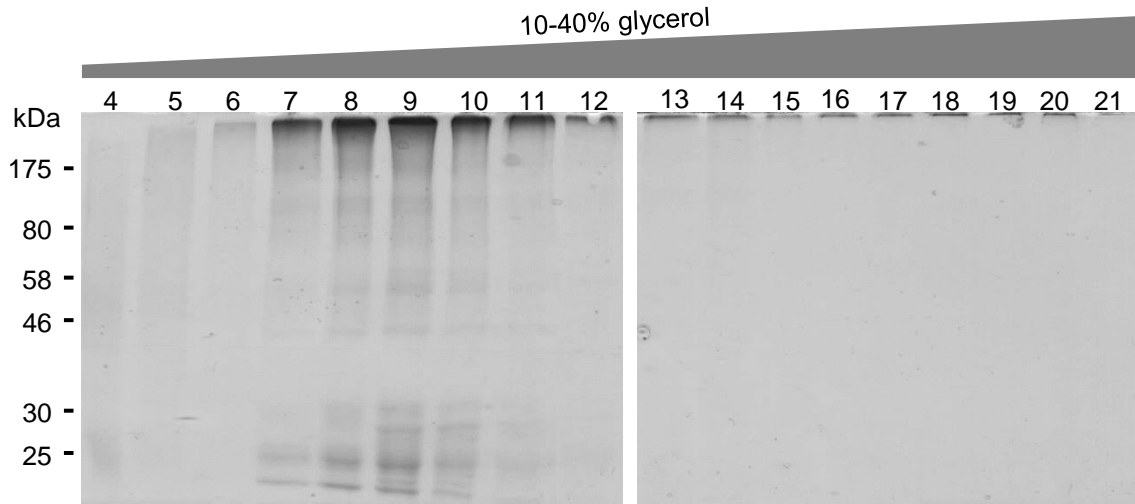


Figure 11: SDS-PAGE analysis of fractions derived from analytical ultracentrifugation of $15S^{GFP}$ complexes cross-linked with 0.8 mM DSS H12/D12 and run on a 10-40% glycerol gradient. Gradient fractions (600 μ l) are indicated by numbers. Fractions 1-3 were not included in the SDS-PAGE analysis.

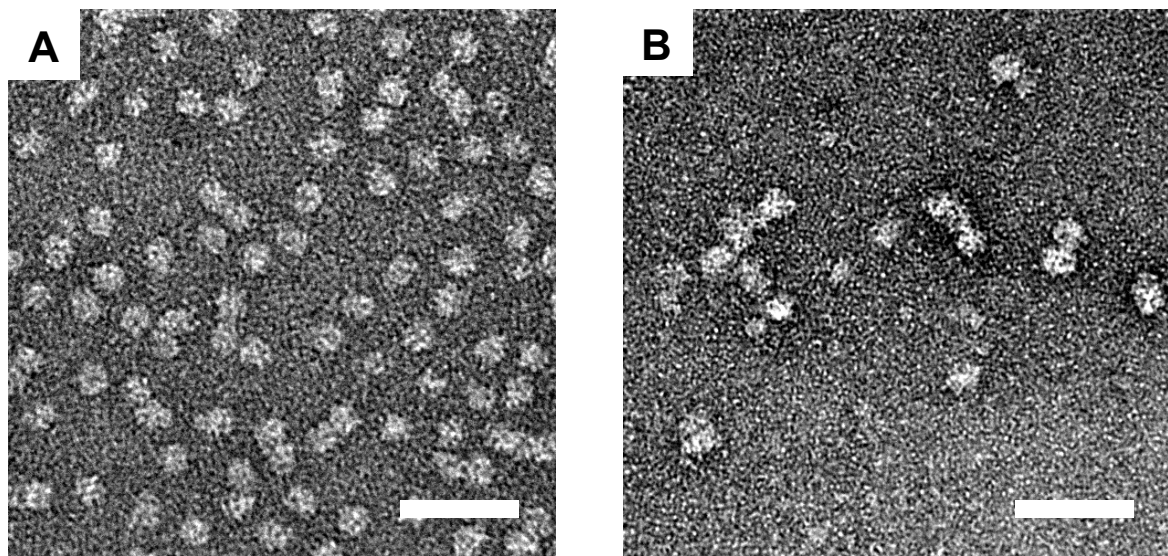


Figure 12: Negatively stained complexes of fraction 9 (A) and fraction 13 (B) of the glycerol gradient displayed in figure 11. Cross-linked $15S^{GFP}$ was stained with 2 % (w/v) Uranyl acetate. The scale bars correspond to a distance of 50 nm.

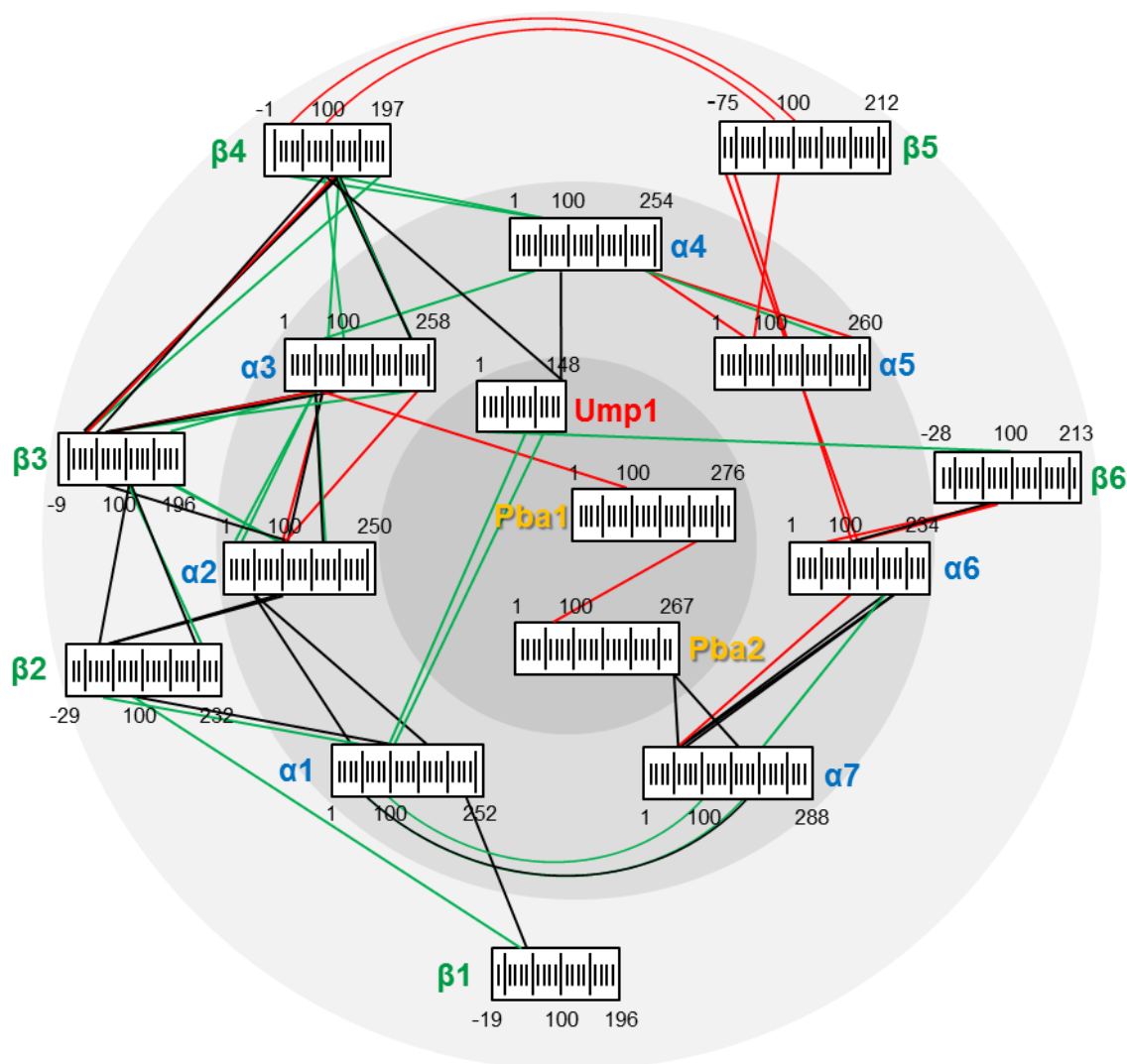
density fractions also contain bands in the upper part of the gel, which are present in fraction 12-21. This indicates that multimeric assemblies of $15S^{GFP}$ complexes were cross-linked to one another (figure 12B).

Analytical ultracentrifugation of cross-linked native 15S complexes failed due to the overall lower yield of this complex compared to $15S^{GFP}$. Enrichment of the 15S complex by this procedure yielded less than 20 μ g of protein, which is below the recommended minimum of

50 µg for subsequent MS analysis. Cross-linking experiments of this complex were therefore performed without additional purification via a glycerol gradient.

5.5 Proteasome precursor complex cross-linking results:

A general overview of all inter-subunit cross-links found in 15S and 15S^{GFP} complexes are given in figure 13 and 14. A detailed description of the cross-linked residues and distances is given in appendix tables 1-4. Appendix tables 5-6 include intra-protein cross-links as well as mono-links, which provide information about the solvent accessibility of lysine residues. All β subunits in these tables have been labeled in a way that the first amino acid is Thr1 or analogues residues (consistent with figure 4). Therefore negative numbers indicate that a residue from pro-peptides or N-terminal extensions has been cross-linked. Both 15S and 15S^{GFP} samples show a cross-linking pattern largely in agreement with the predicted subunit arrangement of a half-proteasome (minus the β 7 subunit), i.e. the order of subunits is the same as in the mature complex. Inter-subunit cross-links overlap in ~41 % of all locations, excluding data found for GFP because it is not present in the native 15S sample. This indicates that both complexes are structurally similar, which is also reflected in the 3D reconstructions from EM data (see section 6.3.1). So while both samples show very similar cross-linking patterns, there are a few minor differences. In 15S^{GFP} no inter-subunit cross-links have been found for β 5 and only one for β 6. In contrast, native 15S displays a larger number of cross-links for both subunits. This suggests a sub-stoichiometric presence of β 5 and β 6 in the 15S^{GFP}, which is also reflected in the peptide mass fingerprinting data for this complex. Possible reasons for the absence of these subunits are described in section 6.4. Heterogeneity concerning the presence or absence of certain subunits was already observed previously when the 15S^{GFP} complex was purified via a Strep-tag on the C-terminus of Ump1-GFP (see figure 4B in Lehmann et al. 2010). The β 5 and β 6 subunits are nevertheless present in sub-stoichiometric amounts as evidenced by both intra- and mono-links found for them (see appendix tables 5-6). The absence of cross-links to Pba1 and the presence of only two cross-links to Pba2 in the 15S^{GFP} complex can be attributed to the fact that the strain used to purify this complex did not overexpress Pba1-Pba2, hence they are underrepresented. Of particular importance are the inter subunits cross- links found for the C-terminus of Ump1 (K144). They overlap in both samples, serving as an internal control to show that Ump1 was correctly incorporated into the 15S^{GFP} complex, despite the additional encumbrance cause by fusing GFP to this end of the chaperone. According to the identified inter-subunit crosslinks, the GFP moiety is located in the central pore between the α and β ring, pointing its N-terminal



43

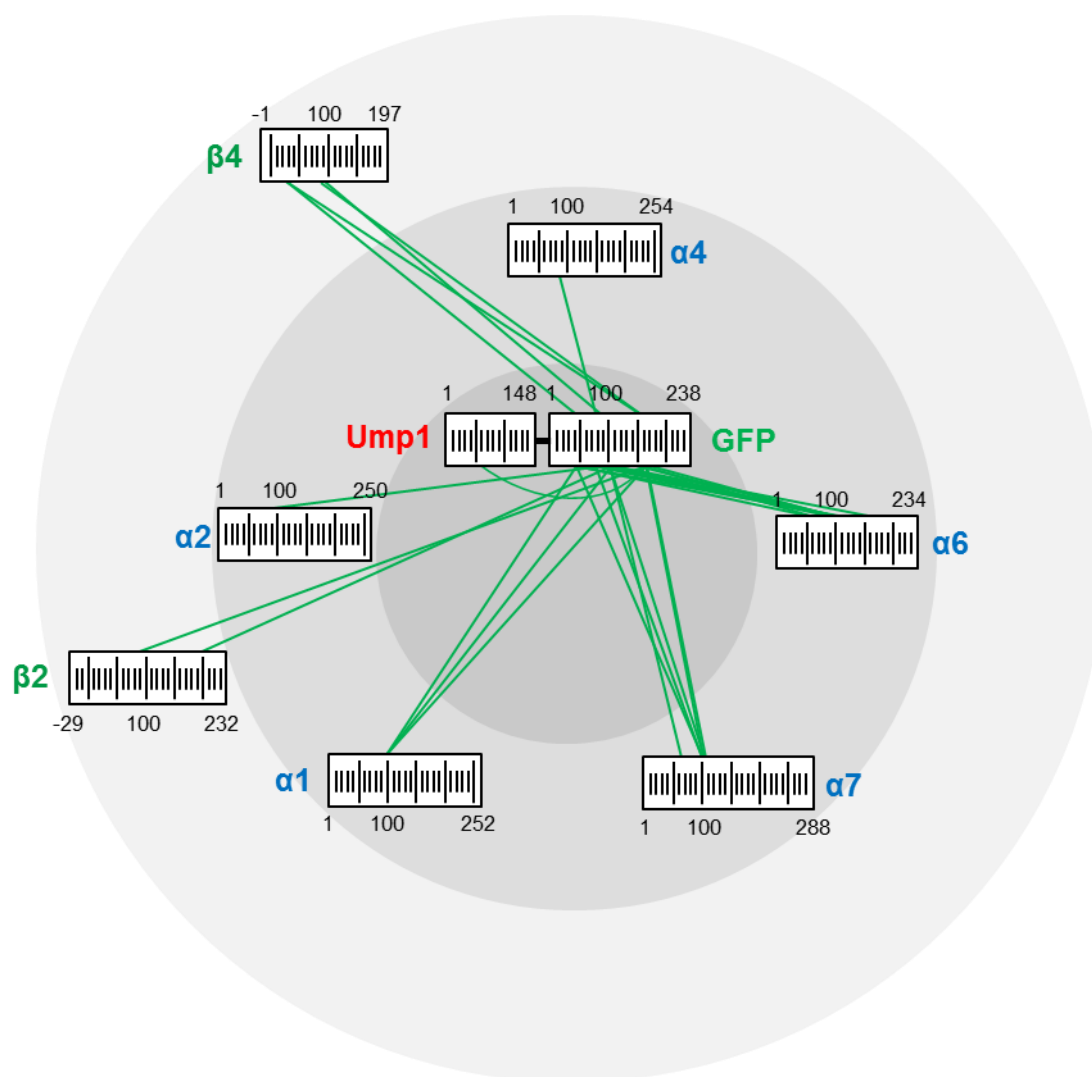


Figure 14: Overview of inter-subunit cross-links found in the $15S^{GFP}$ complex using cross-linking and MS analysis. Only cross-links that relate to the GFP moiety fused to the C-terminus of Ump1 are shown here. Subunits are displayed the same as in figure 13. Subunits that did not cross-link to GFP have been omitted from this chart.

Both samples, $15S$ and $15S^{GFP}$ complexes, feature a small number of inter-subunit cross-links that are not consistent with the predicted subunit arrangement in the proteasome precursor, e.g. they were found between non neighboring subunits or span a distance that exceeds the combined length of the cross-linker and two lysine side chains (~ 30 Å). These cross-links are summarized in appendix table 3. They can probably be attributed to transient dimers of $15S$ complexes that can be found in all samples but which can be detected more frequently after cross-linking the complexes. Figure 15 shows the micrograph of a cross-linked sample containing multiple dimers. While the overall number of dimeric $15S$ complexes is small compared to monomeric $15S$ complexes, they can nevertheless be the source of aberrant cross-links because they bring subunits into close contact that are normally not neighbors. The

reason why some subunits (in particular $\alpha 6$) are over-represented in the aberrant cross-links is puzzling and suggests that these dimers are not entirely random. Whether there is any biological significance to this issue remains to be investigated.

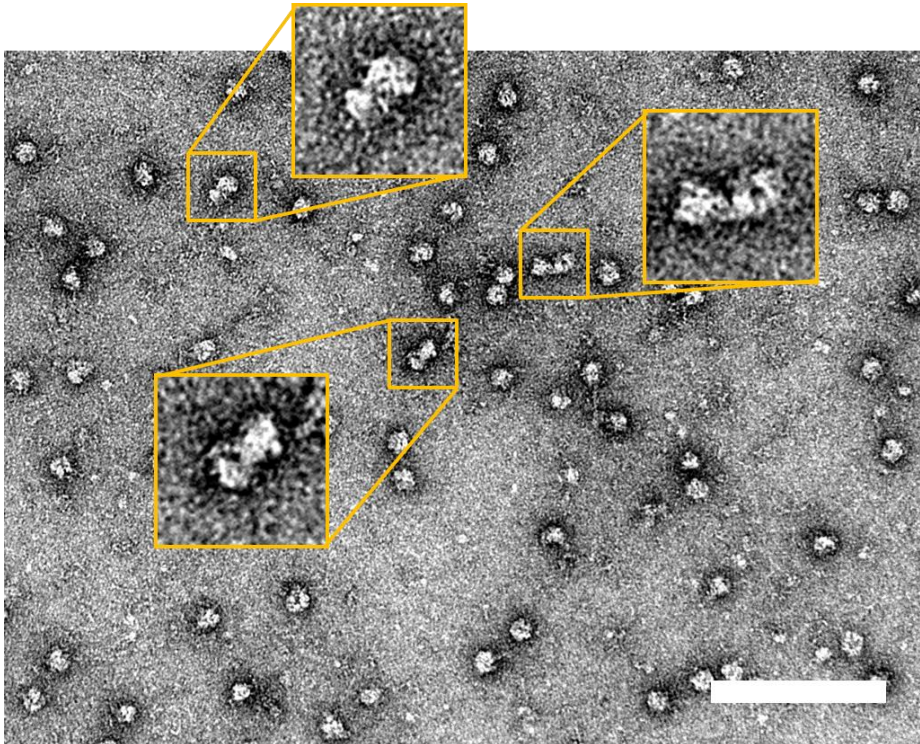


Figure 15: 15S complexes cross-linked with 0.8 mM DSS H12/D12 and stained with 2% (w/v) Uranyl acetate. Most 15S appear as monomeric complexes but the cross-linking reaction has also produced 15S dimers (orange encircled areas) that can be the cause of aberrant cross-links. The scale bar corresponds to a distance of 100 nm.

6. EM results:

6.1 EM analysis of purified protein complexes:

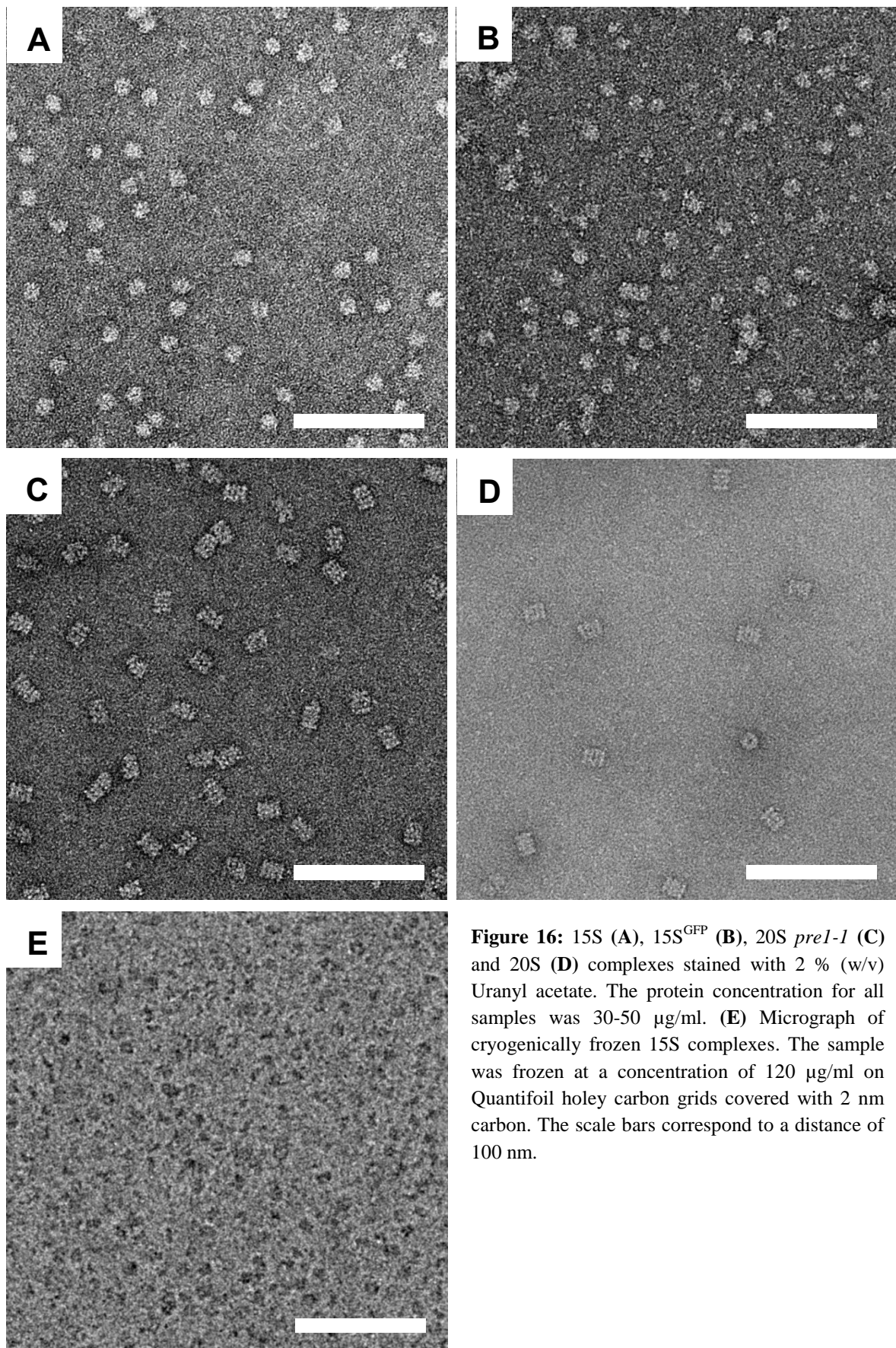
Figure 16 shows affinity purified and negatively stained 15S, 15S^{GFP} 20S *prel-1* and 20S complexes on micrographs collected for use in image analysis on the FEI G2 Spirit. 20S *prel-1* complexes show a low occupancy of Pba1-Pba2 heterodimers with only ~10-15 % of the complexes exhibiting additional density on either end. While the native 15S sample generally shows less aggregation and heterogeneity than the one containing GFP fusion proteins, all were judged to be of sufficient quality for image collection and analysis.

6.2 Classification of EM data:

6.2.1 Eigenimage analysis:

Eigenimages derived from multivariate statistical analysis (MSA) of negatively stained 15S and 15S^{GFP} complexes using IMAGIC are shown in figure 17. These images were used to explore symmetry parameters of both complexes. Of the 69 eigenvectors used in each classification, only the most significant images corresponding to eigenvectors 1 to maximally 8 are displayed here. The first eigenimage in each row shows the total sum of all particles that were analyzed. The 5th and 6th eigenimage of the 15S complexes show 7-fold rotational symmetry with an apparent break (figure 17A). This indicates the pseudo-heptameric symmetry of the 15S α and β rings. The 4th eigenimage appears to show an 8-fold rotational symmetry. This is due to the large number of tilted particles present in the dataset and not due to an actual 8-fold symmetry. Obvious side views were excluded from the eigenimage analysis, as they were not expected to show any pseudo symmetry. Eigenimages obtained from 15S^{GFP} data resemble those of the 15S complex MSA, showing rotational 7-fold symmetry along the pore axis (figure 17B).

To identify extra density evoked by the GFP moiety, particles corresponding to top and bottom views of 15S and 15S^{GFP} complexes (i.e. $\beta = 0-25^\circ$ and $\beta = 155-180^\circ$) were extracted, combined and investigated by MSA. Figure 17C shows eigenimages gained from this analysis, with the 2nd to 7th eigenimage displaying the expected 7- and 8-fold rotational symmetry already observed before. In addition, the 4th to 8th image show a difference in the middle of the complex. This difference is especially pronounced in the 8th image. Since GFP fused to the C-terminus of Ump1 is the only additional protein introduced into this complex,



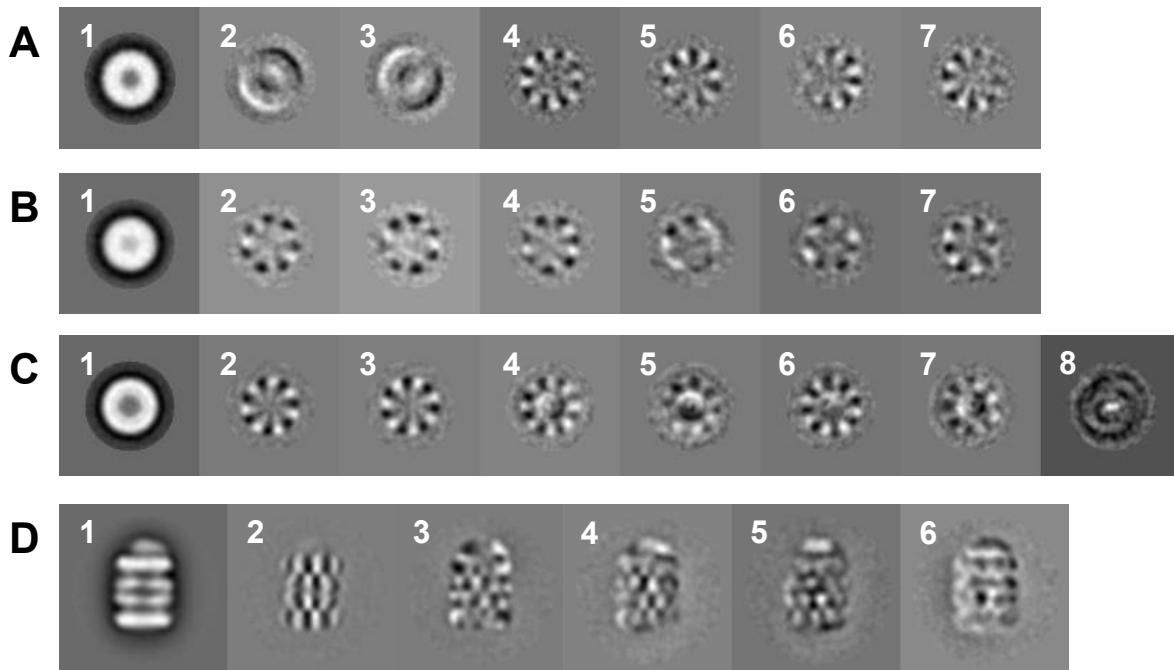


Figure 17: Eigenimages generated from negatively stain proteasome precursor complexes. **(A)** The first 7 eigenimages generated from 15S data. **(B)** The first 7 eigenimages generated from 15S^{GFP} data. **(C)** Eigenimages showing the difference between end views of 15S and 15S^{GFP} complexes. **(D)** The first 6 eigenimages generated from side views of 20S *preI-1* data analysis.

the difference in the middle of the rings may correspond to GFP (see also section 5.5 and 6.4 for more information).

In the case of the 20S *preI-1* complex, eigenimage analysis was performed to probe particle images for the presence or absence of Pba1-Pba2 densities. The eigenimages were also screened for differences between particles related to different stages of maturation, in order to explore possible ways to improve sorting and increase the homogeneity of the dataset. End views of the complex were omitted from the analysis as they were not expected to contain useful information about the subunit arrangement that cannot be better visualized in side views. The 2nd to 5th eigenimage (figure 17D) shows differences in the α and β subunits caused by the rotation of the complex around the (vertical) pore axis, which results in the alternating subunit pattern observed in these images. The 3rd to 6th eigenimage shows differences originating from the Pba1-Pba2 density at the end of the proteasome. While these differences may be related to the presence or absence of the chaperones in particle images, several of the eigenimages (in particular the 3rd and 4th) actually seem to relate to a rotation around the pore axis which causes the chaperones to appear in different spots atop the barrel-like 20S density. The 6th eigenimage appears to relate to images containing elongated

particles. However this difference was found to relate to slight misalignments of particles along the vertical axis. In summary, none of the eigenvectors seems to relate to differences caused by working with particles in different stages of maturation. Heterogeneity due to the presence or absence of Pba1-Pba2 was also kept at a minimum.

6.2.2 Class average analysis:

Representative class averages derived from MSA of negatively stained and cryogenically frozen 15S complexes are shown in figure 18A and 18B respectively. Class averages from both datasets show round end views (image 1-3) and a double-layered side view (image 4-6). End view class averages show additional density in the middle of the rings, which is expected to correspond to the Pba1-Pba2 heterodimer. However, class averages representing side views do not show much density at the end of the rings. The putative Pba1-Pba2 density is particularly difficult to identify in cryo-EM data (figure 18B image 4-6). This is in stark contrast to the class averages obtained from 20S *pre1-1* negative stain data (figure 18C). Here the Pba1-Pba2 heterodimer forms very pronounced densities on one or both ends of the mutated 20S (see images 1-4) which are clearly separated from the rings they are located on.

While cryo-EM data of the 15S complex collected on the Tecnai G2 Spirit still shows features also found in negative stain data, the overall quality of the classification is worse. This is probably owed to the reduced signal-to-noise ratio of cryo-EM data which makes the alignment of images of a small complex such as the 15S more difficult. This problem was exacerbated in cryo-EM data collected at the Tecnai F20 (not shown). Data from this microscope was classified but the class averages did not resemble the shapes found in negative stain and cryo-EM data from the Tecnai G2 Spirit. This problem could not be solved by increasing the particle count per class to improve the signal-to-noise ratio. Possible reasons include the weaker signal-to-noise ratio caused by collecting data at a higher extraction voltage (200 kV on the Tecnai F20 versus 120 kV on the Tecnai G2 Spirit). Analysis of data collected at the Tecnai F20 was not pursued further, as this problem was expected to severely decrease the quality of 3D reconstructions instead of improving the resolution as originally intended.

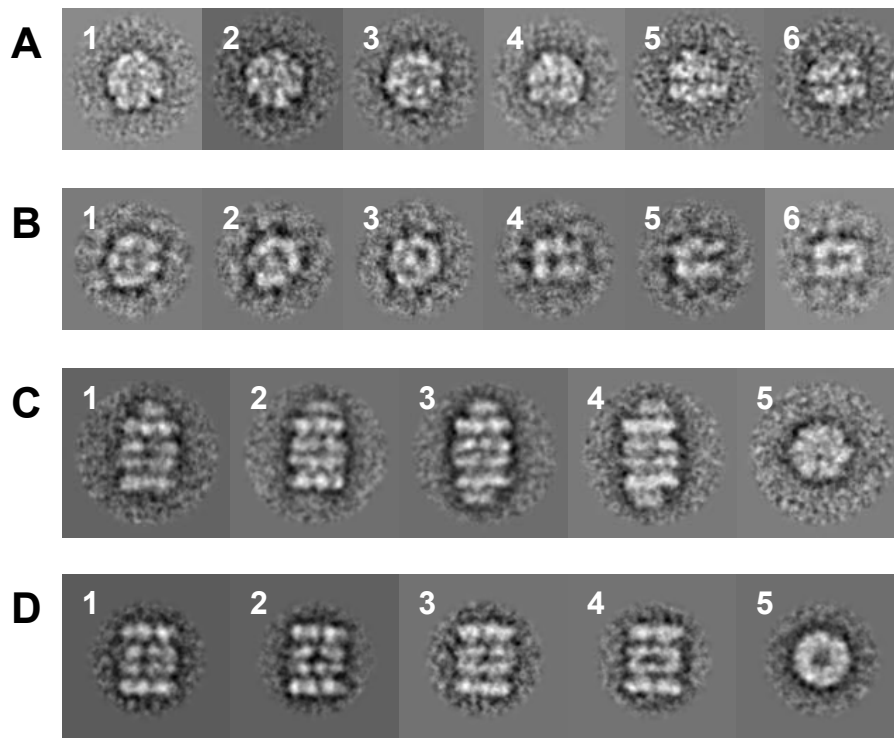


Figure 18: Class averages of particles generated from 3 rounds of combined MSA and MRA. **(A)** Negatively stained 15S complexes. Putative top views are shown in image 1-3, while side views are shown in image 4-6. **(B)** cryogenically frozen 15S data from the Tecnai G2 Spirit. Putative top views are shown in image 1-3, while side views are shown in image 4-6. **(C)** Negatively stained 20S *preI-1* complexes. Images 1-2 show side views of 20S *preI-1* with a single Pba1-Pba2 heterodimer attached, while image 3-4 show two heterodimers attached to either end. Image 5 shows an end view of the complex. **(D)** Negatively stained 20S complexes. Images 1-4 show side views of 20S, while image 5 shows an end view of the complex.

6.3 Refinement:

6.3.1 15S and 15S^{GFP} 3D reconstructions:

The initial 15S three dimensional model obtained by angular reconstitution in IMAGIC was refined using a total of 10156 negatively stained particles and projection matching in SPIDER. The resulting 3D map of the 15S complex served as a template for reconstruction of the 15S^{GFP} complex using 3873 particles. Figure 19 and figure 20 show the final 3D reconstructions, as well as class averages and corresponding re-projections in the Euler angle directions assigned to the class averages. In both reconstructions, the projections match their corresponding re-projections reasonably well, hence the refinement was assumed to be reliable.

The 3D reconstruction of the 15S^{GFP} complex resembles that of the native 15S complex with few differences. The densities of Pba1 and Pba2 situated in the ring of α subunits are not as defined in the 15S^{GFP} complex. Densities that constitute the two rings slightly change in size compared to the native 15S, but the core densities are still in the same place, as seen in figure 21. These differences arise from a higher degree of heterogeneity in the purified 15S^{GFP} complex in comparison with the 15S complex, which in turn can influence the alignment. The heterogeneity is also reflected in the micrograph shown in figure 16B.

The most intriguing difference between the 3D reconstructions of 15S and 15S^{GFP} complexes is the extra density inside the pore between the two rings (green encircled area in figure 20B) when the maps are rendered at thresholds corresponding to their respective molecular weight of 432 and 460 kDa. Cross-linking data and eigenimage analysis predict this to be the location of GFP from the Ump1-GFP fusion protein. The ability of the proteasome antechambers to hold entire GFP molecules has been demonstrated before (Sharon et al. 2006), although GFP was unfolded in this case. However the extra density found in the EM map of the 15S^{GFP} complex is too small for GFP (see figure 29). The fact that this density is so small even though GFP is predicted to reside inside the pore might be owed to the GFP being flexibly attached. This is described in more detail in section 6.4.

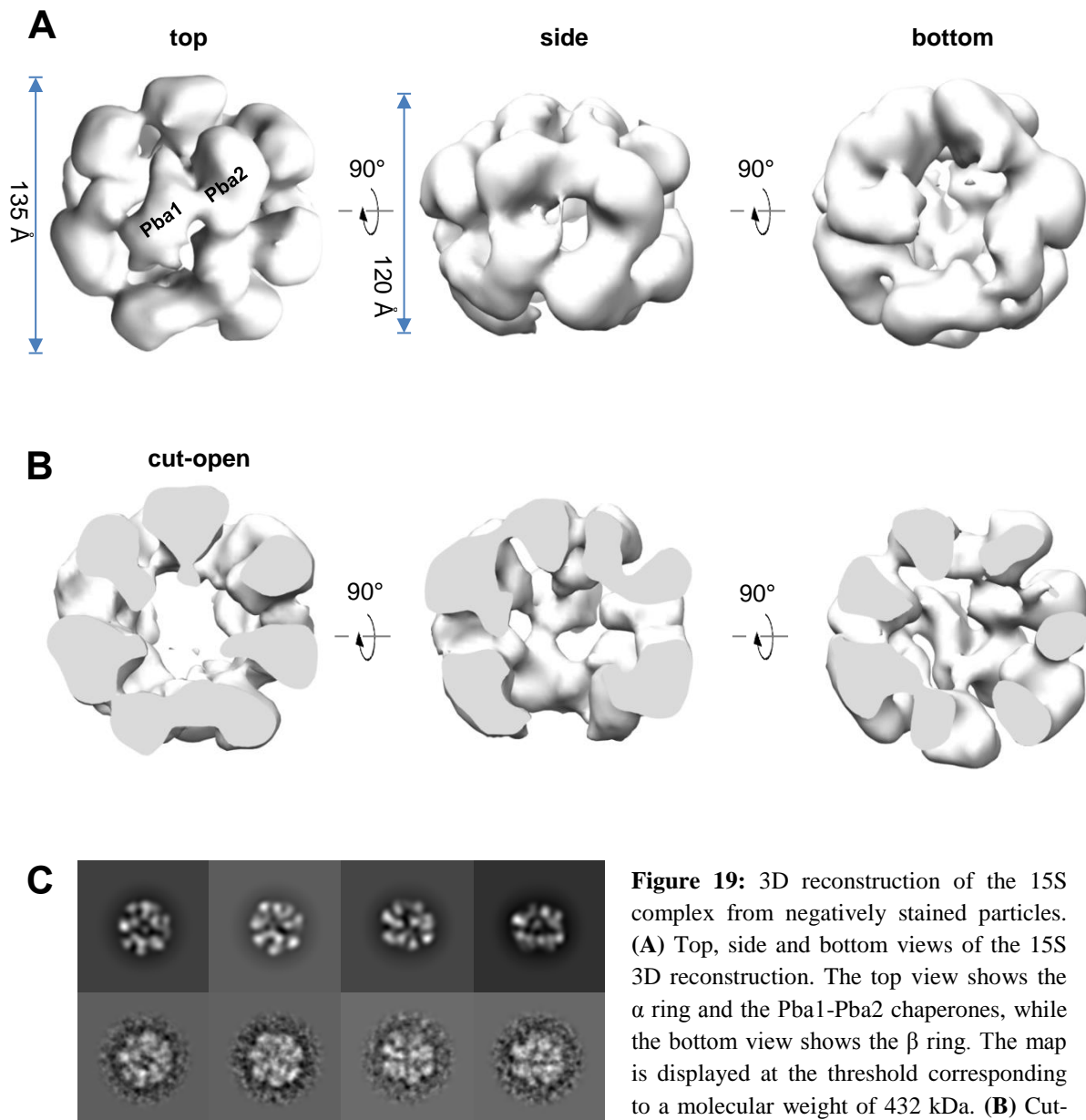


Figure 19: 3D reconstruction of the 15S complex from negatively stained particles. **(A)** Top, side and bottom views of the 15S 3D reconstruction. The top view shows the α ring and the Pba1-Pba2 chaperones, while the bottom view shows the β ring. The map is displayed at the threshold corresponding to a molecular weight of 432 kDa. **(B)** Cut-open views of the 3D reconstruction displayed in the same orientation as in (A). **(C)** Selected class averages of the MSA treated dataset of the 15S complex (bottom row) and re-projections (top row) of the final 15S 3D reconstruction in the Euler angle directions assigned to the class averages. Each class average consists of ~10 particles/class.

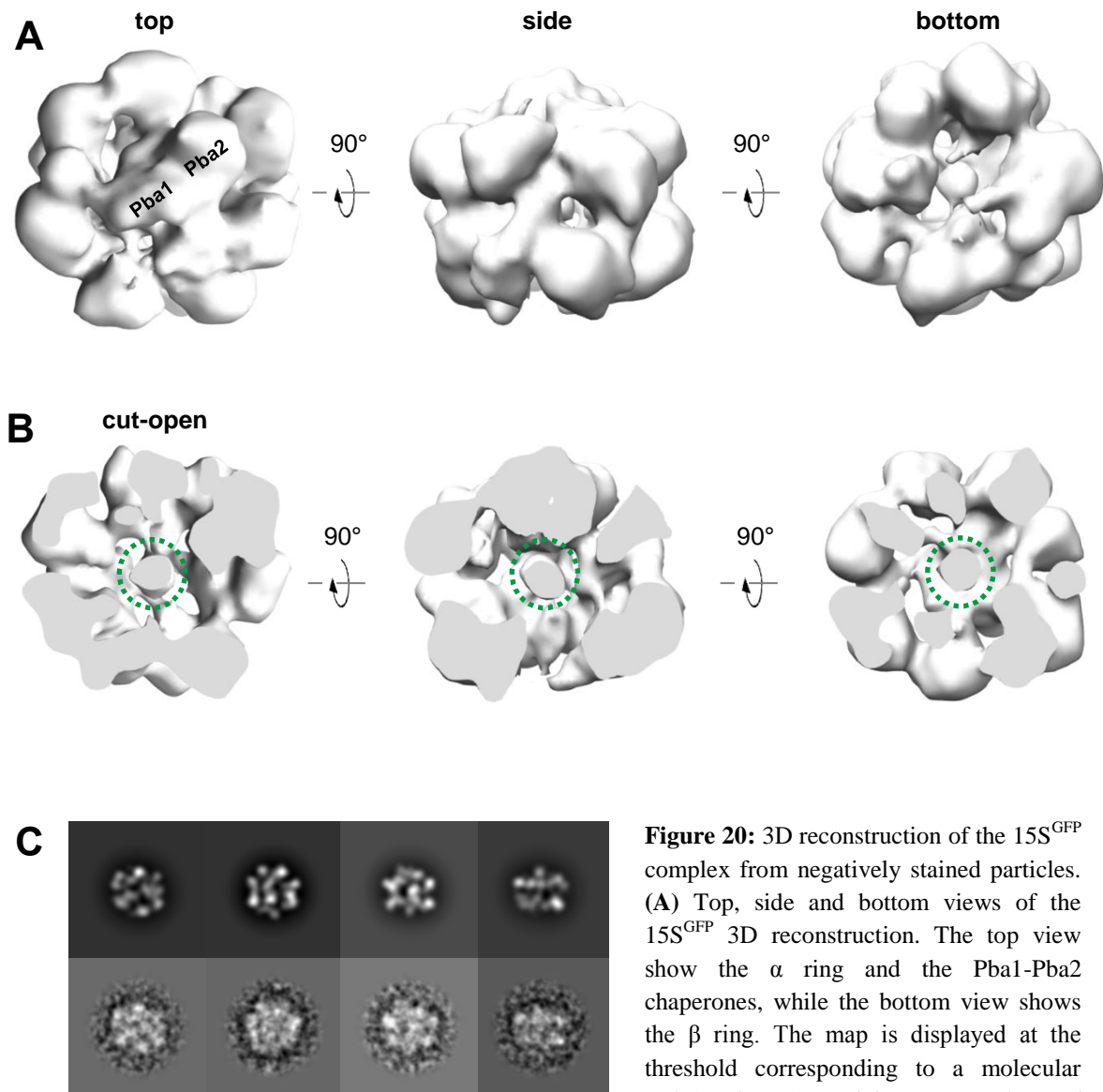


Figure 20: 3D reconstruction of the 15S^{GFP} complex from negatively stained particles. (A) Top, side and bottom views of the 15S^{GFP} 3D reconstruction. The top view shows the α ring and the Pba1-Pba2 chaperones, while the bottom view shows the β ring. The map is displayed at the threshold corresponding to a molecular weight of 460 kDa. (B) Cut-open views of the 3D reconstruction displayed in the same orientation as in (A). The green encircled areas show the putative location of the GFP moiety as suggested by cross-linking data and the EM density. (C) Selected class averages of the MSA treated dataset of the 15S^{GFP} complex (bottom row) and re-projections (top row) of the final 15S^{GFP} 3D reconstruction in the Euler angle directions assigned to the class averages. Each class average consists of ~10 particles/class.

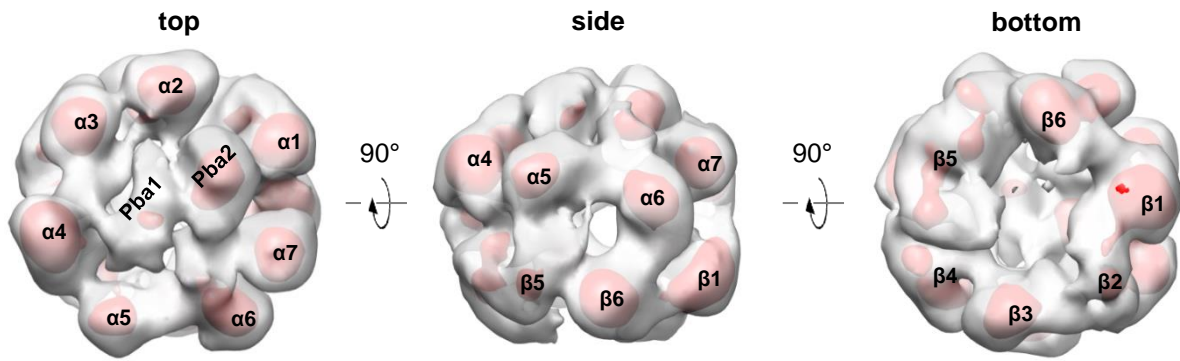


Figure 21: Comparisons between 3D reconstructions of 15S and 15S^{GFP} complexes. The 15S reconstruction is shown in transparent grey and 15S^{GFP} core densities in red.

6.3.2 20S *pre1-1* and 20S 3D reconstructions:

The initial 20S *pre1-1* three-dimensional model obtained by angular reconstitution in IMAGIC was refined using a total of 12609 negatively stained particles and projection matching in SPIDER. To avoid reference bias an initial three-dimensional model was constructed for the mature 20S as well, instead of relying on a low pass filtered crystal structure. This model was refined using a total of 4445 negatively stained particles. Figure 22 and 23 show the final 3D reconstructions of 20S *pre1-1* and 20S complexes as well as class averages and corresponding re-projections in the Euler angle directions assigned to the class averages. For both reconstructions, the projections match their corresponding re-projections, hence the refinement was assumed to be reliable. Apart from the additional density on top of the 20S *pre1-1* complex that can be attributed to the chaperones Pba1 and Pba2, there are additional differences between the late precursor and the mature complex. Those are described in more detail below in section 6.5.

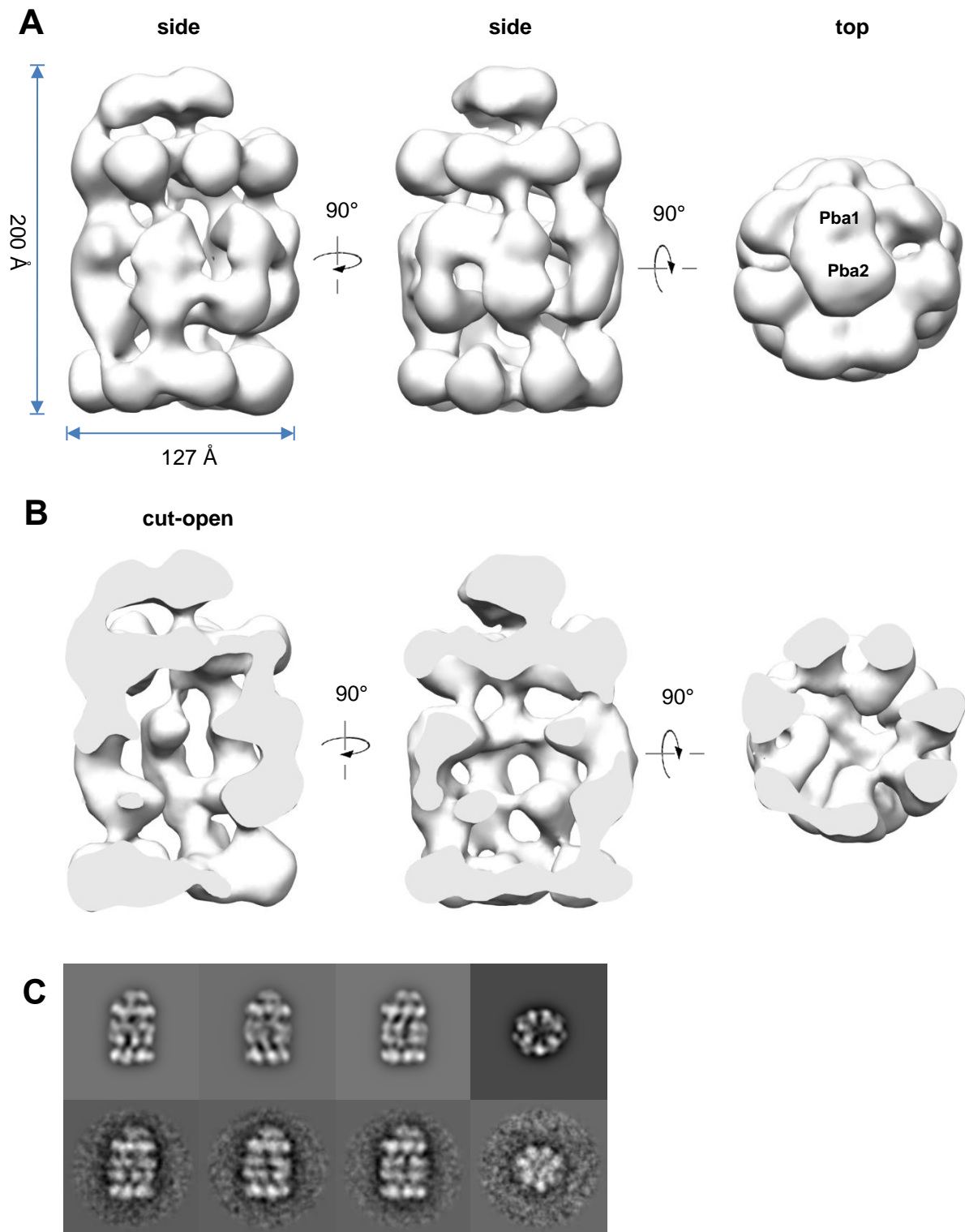


Figure 22: 3D reconstruction of the 20S *pre1-1* complex from negatively stained particles. **(A)** Two side views and a top view of the 20S *pre1-1* reconstruction. Side views show the stack-like $\alpha 7$ - $\beta 7$ - $\beta 7$ - $\alpha 7$ arrangement with the Pba1-Pba2 density on top. The top views show the Pba1-Pba2 chaperones in the foreground with the α ring beneath. **(B)** Cut-open views of the 3D reconstruction displayed in the same orientation as in (A). **(C)** Selected class averages of the MSA treated dataset of the 20S *pre1-1* complex (bottom row) and re-projections (top row) of the final 20S *pre1-1* 3D reconstruction in the Euler angle directions assigned to the class averages. Each class average consists of ~10 particles/class.

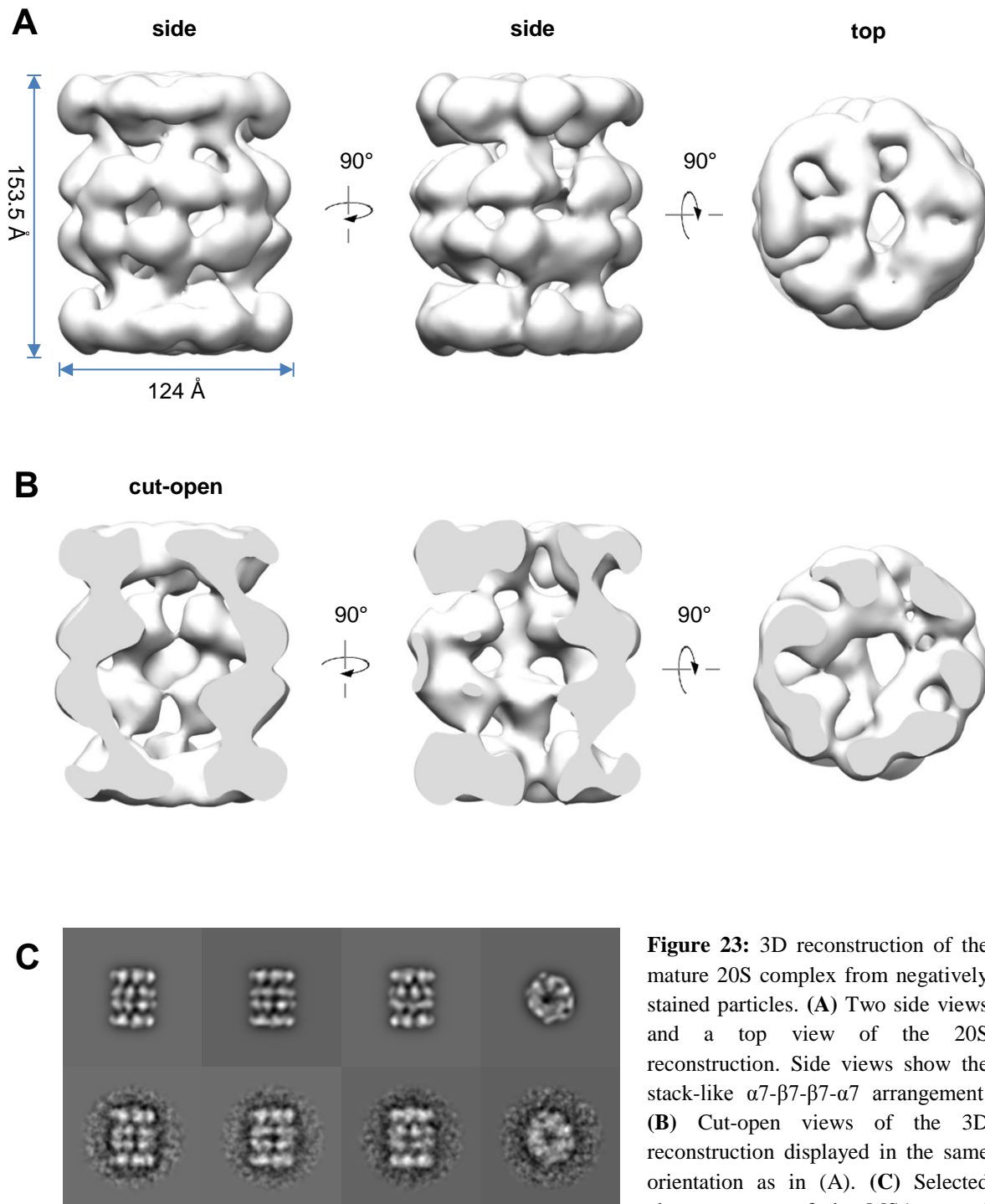
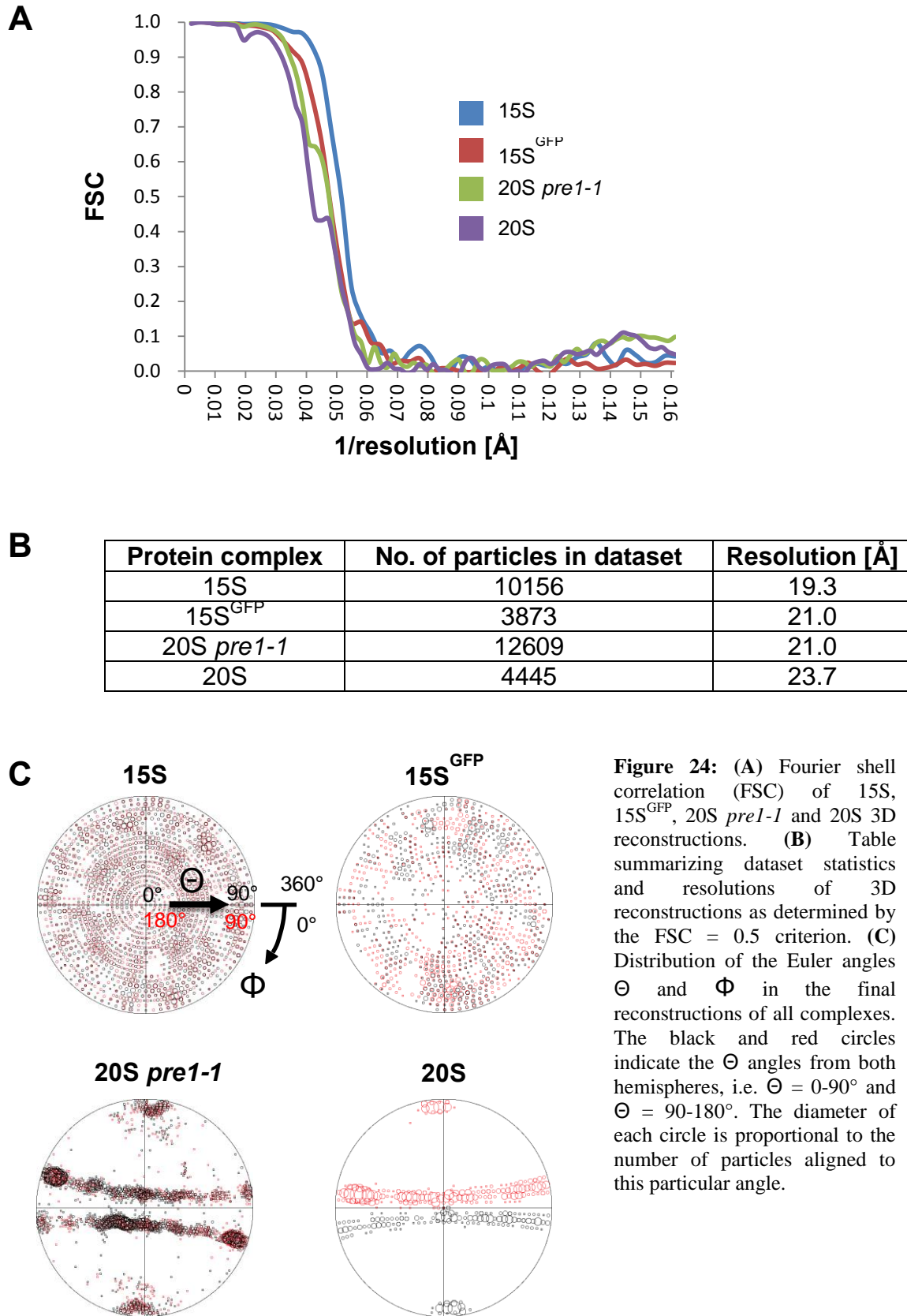


Figure 23: 3D reconstruction of the mature 20S complex from negatively stained particles. **(A)** Two side views and a top view of the 20S reconstruction. Side views show the stack-like $\alpha 7$ - $\beta 7$ - $\beta 7$ - $\alpha 7$ arrangement. **(B)** Cut-open views of the 3D reconstruction displayed in the same orientation as in (A). **(C)** Selected class averages of the MSA treated dataset of the 20S complex (bottom row) and re-projections (top row) of the final 20S 3D reconstruction in the Euler angle directions assigned to the class averages. Each class average consists of ~10 particles/class.



6.3.3 Resolution and dataset statistics:

The resolutions of 3D reconstructions of 15S, 15S^{GFP}, 20S *prel-1* and mature 20S complexes were determined by plotting the Fourier shell correlation (FSC) against the reciprocal resolution value as shown in figure 24A using the FSC 0.5 criterion. 3D reconstructions were not masked for determining the resolution. All 3D reconstructions show a similar resolution close to 20 Å, which corresponds to the Uranyl acetate grain size. Figure 24C shows the Euler angle distribution of all the particles used for 3D reconstruction of the -15S, -15S^{GFP}, -20S *prel-1* and mature 20S complexes after refinement with projection matching. While certain angles are present more often than others, there is no clear preferential orientation of the 15S and 15S^{GFP} complexes on the grid. The reconstructions of the 20S *prel-1* and mature 20S complexes show a preferential orientation on the grid, generating predominantly side and end views of the particle. This is owed to the elongated shape of the particle, which discourages other orientations.

6.3.4 Determination of handedness - 15S complex:

In order to unambiguously assign the hand to the 3D reconstruction of the 15S complex the program CTFTILT (Mindell & Grigorieff 2003) was used. Therefore pairs of micrographs with a tilt angle difference of 15° were collected on the Tecnai G2 Spirit using negatively stained complexes. The Pex1/6 AAA+ ATPase complex assembled in the presence of ATPγS (Ciniawsky et al.) was used as a control because the handedness can be easily deduced from the reconstruction by fitting crystal structure homology models based on the AAA+ ATPase p97. While the control experiment still contains some outliers, the majority of particles were aligned in a way that the difference between the right and the wrong handedness can be determined. Outliers are defined as particles that are not tilted in-plane, which results in a large out-of-plane error. Alignment output parameters of the Pex1/6 complex cluster in a tilt angle range of ~10-18°, which is an acceptable approximation of the real tilt angle of 15° (figure 25A). Particles aligned to the model with the correct handedness were assigned output parameters that cluster on the tilt axis of the goniometer (as they should), while those aligned to the model with the wrong handedness are clearly clustering off the axis. In summary, the handedness determination worked for the Pex1/6 control sample. The alignment output parameters of the 15S sample however, are scattered with a high percentage of out-of-plane tilt angles (figure 25B). Neither the tilt angle nor the handedness can be deduced from the data, as the particles do not show consistent alignment parameters close to or on the tilt axis. A possible explanation is that the 15S complex does not possess structural features that could

give away the handedness easily. Side views of the complex look very similar to each other. The top views show a slightly asymmetric arrangement of the Pba1-Pba2 heterodimer but it is not as pronounced as in the 20S *pre1-1* complex and therefore probably too small to help with the determination of the handedness. For this reason it was decided that the tilt pair method is not suitable to determine the handedness of the 15S 3D reconstruction. While examples exist for protein complexes of comparable or slightly higher molecular weight than 15S on which the tilt pair method worked (Henderson et al. 2011), said examples consist of completely asymmetric complexes that give away the handedness much more easily.

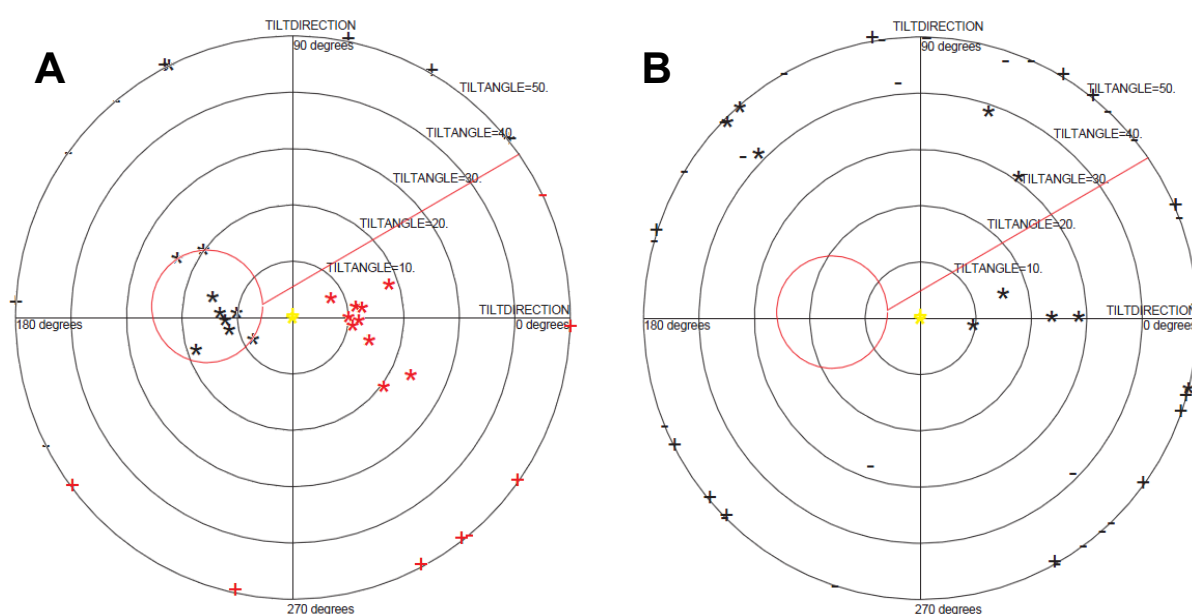


Figure 25: Plots of alignment output parameters of the handedness determination showing the tilt angle and tilt direction of individual particles. **(A)** Pex1/6 control sample. Particles aligned to the correct hand are shown in black, those aligned to the wrong hand in red. **(B)** 15S particles aligned to one of the two possible hands are shown in black. The red diagonal line in both plots indicates the tilt axis of the goniometer, while the area inside the red circle represents the space in which alignment output parameters have to cluster if the particles were aligned to the correct hand. Black concentric rings provide a measure for the tilt angle found for each individual particle. Alignment parameters of particles marked “*” were found to contain Euler angles consistent with in-plane tilts of specimen. Particles marked “-” are outliers and exceed the average out-of-plane error while those marked “+” show a deviation from this average by a factor of at least 1.5.

Instead of using tilt pair analysis, the 15S complex handedness was determined by fitting crystal structures of proteasome subunits using restraints from both the EM map and cross-linking data. The EM densities found for Pba1 and Pba2 are of particular importance, as they narrow down the number of possible orientations in an otherwise pseudo-symmetrical complex. Using a process of elimination, this lead to the crystal structure fit shown in section

6.4, figure 26. This fit accommodates the EM map reasonably well, in addition to almost every cross-link formed between neighbouring subunits in the 15S and 15S^{GFP} complexes (listed in appendix table 1-2). Only a single inter-subunit cross-link between neighbouring subunits was violated by this fit. This cross-link is described in section 6.4. Every other crystal structure fit either cannot accommodate the EM map reasonably well or it creates large spaces between individual subunits that violate multiple cross-links simultaneously. In summary, the combined use of restraint from EM and cross-linking data approximates the solution to the handedness problem, although the latter could not be solved with absolute certainty due to the failure of the tilt pair method.

6.3.5 Determination of handedness - 20S *pre1-1* complex:

As the shape of the 20S *pre1-1* 3D reconstruction resembles that of the 20S-Pba1-Pba2 crystal structure (Stadtmueller et al. 2012), the handedness of the 20S *pre1-1* complex could be confirmed without resorting to tilt pair analysis. The low pass filtered map of the crystal structure was fitted into the EM density and cross correlation coefficients were calculated using CHIMERA for either hand of the complex (0.6278 vs. 0.6259). The hand with the higher score was assumed to be the correct one. A similar procedure was performed for the 3D reconstruction of the 20S complex to determine the handedness and subunit register, calculating cross correlations between a low pass filtered 20S crystal structure (Groll et al. 2000) and the EM map.

6.4 Subunit fits for negative stain reconstructions - 15S and 15S^{GFP} complexes:

The 3D reconstruction of the 15S model from negatively stained particles shows the expected double-layered structure consisting of one α and one β ring. Subunits were fitted into this map taking both the electron density and cross-linking data into account (see figure 26). Fits were optimized in a way that showed the smallest deviation from the positions of subunits in the crystal structure of the mature 20S complex, as cross-linking data does not support a model that changes the subunit register through large shifts of individual subunits. EM and cross-linking data point to a number of features, which clearly distinguishes the 3D reconstruction from a half-proteasome as one would expect from crystal structures. First of all the α ring has an increased diameter of 11 Å compared to the mature complex (see figure 19 and 23 for details). All α subunits are moved away from the center of the ring in a radial fashion,

increasing the lateral distance between neighboring subunits. The pore of the α ring harbors two distinct densities, which correspond to the Pba1-Pba2 heterodimer. Both Pba densities are partially submerged in the α ring pore. This arrangement positions them almost in the middle of the ring, being slightly more oriented toward $\alpha 1$ and $\alpha 7$. The height of the α and β ring alone does not differ significantly from the mature complex.

While the α subunits maintain the appearance of an ordered structure despite their increased distance to each other, the β subunits do not resemble a well-ordered ring. Apart from the fact that the $\beta 7$ subunit is missing from the 15S complex, the rest of the subunits still manage to partially “close the ring” by increasing the distances between each other in certain places (figure 27). This is most prominently seen for $\beta 1$ and $\beta 6$, which are both moved in the direction of the area where $\beta 7$ is located in the mature complex by 15 Å and 12 Å respectively. Similarly the subunit $\beta 5$ is moved in the same direction as $\beta 6$ by 3 Å, away from $\beta 4$. The latter is moved away from the center of the β ring by 8.5 Å and into the direction of the subunits $\alpha 3$ and $\alpha 4$ by 5 Å, which leads to a position between the two rings. By positioning the $\beta 4$ subunit closer to the α ring, two cross-links ($\alpha 4$ K66 to $\beta 4$ K28 and $\beta 5$ K71 to $\beta 4$ K89) can be accommodated which would be violated by the subunit arrangement of the mature 20S. This supports the hypothesis that $\beta 4$ does indeed occupy this position in the precursor complex and shifts its position during the maturation process.

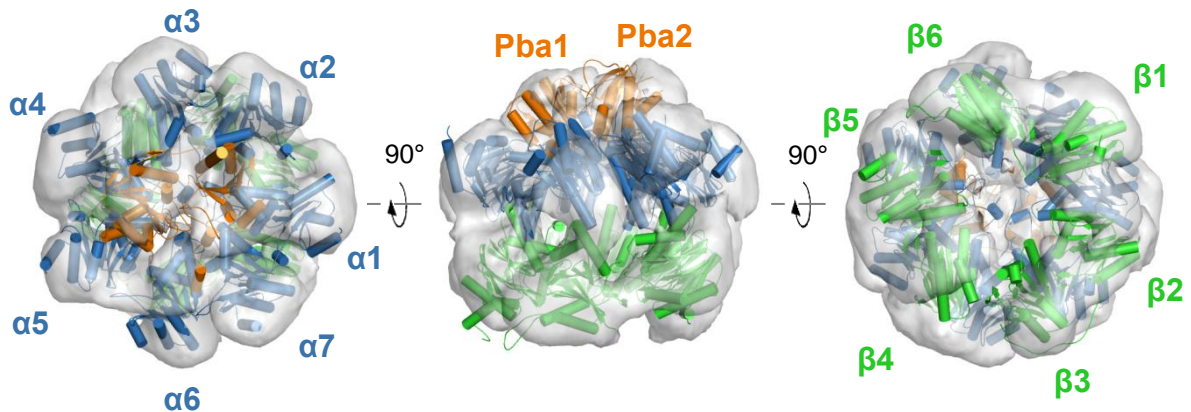


Figure 26: Fitting crystal structure subunits into the 15S complex EM density. Subunits were fitted individually using restraints imposed by the EM density and cross-linking data. α subunits are coloured blue, β subunits green and Pba chaperones orange. The crystal structures of proteasome subunits have been extracted from the *S. cerevisiae* 20S proteasome (Groll et al. 2000).

Density between the subunits $\beta 5$ and $\beta 4$ extends out of the ring and toward the middle of the pore, but the largest part of the pore remains open. A similar but smaller density is observed between the subunits $\beta 1$ and $\beta 2$. It is assumed that these extra densities (that cannot be accommodated by any of the fitted crystal structures) correspond to the pro-peptides of the

catalytic subunits $\beta 1$, $\beta 2$ and $\beta 5$ or parts thereof. This is due to their proximity to these subunits and the fact that in the next maturation step they would be located at the interface between two half-proteasomes, where pro-peptides play a role during dimerization (Li et al. 2007). This implies that the pro-peptides are at least partially structured. Due to the observation that some subunits in the β ring are arranged differently compared to the mature 20S, two 15S complexes would not be able to dimerize prematurely, as binding partners from the opposite β ring have shifted positions.

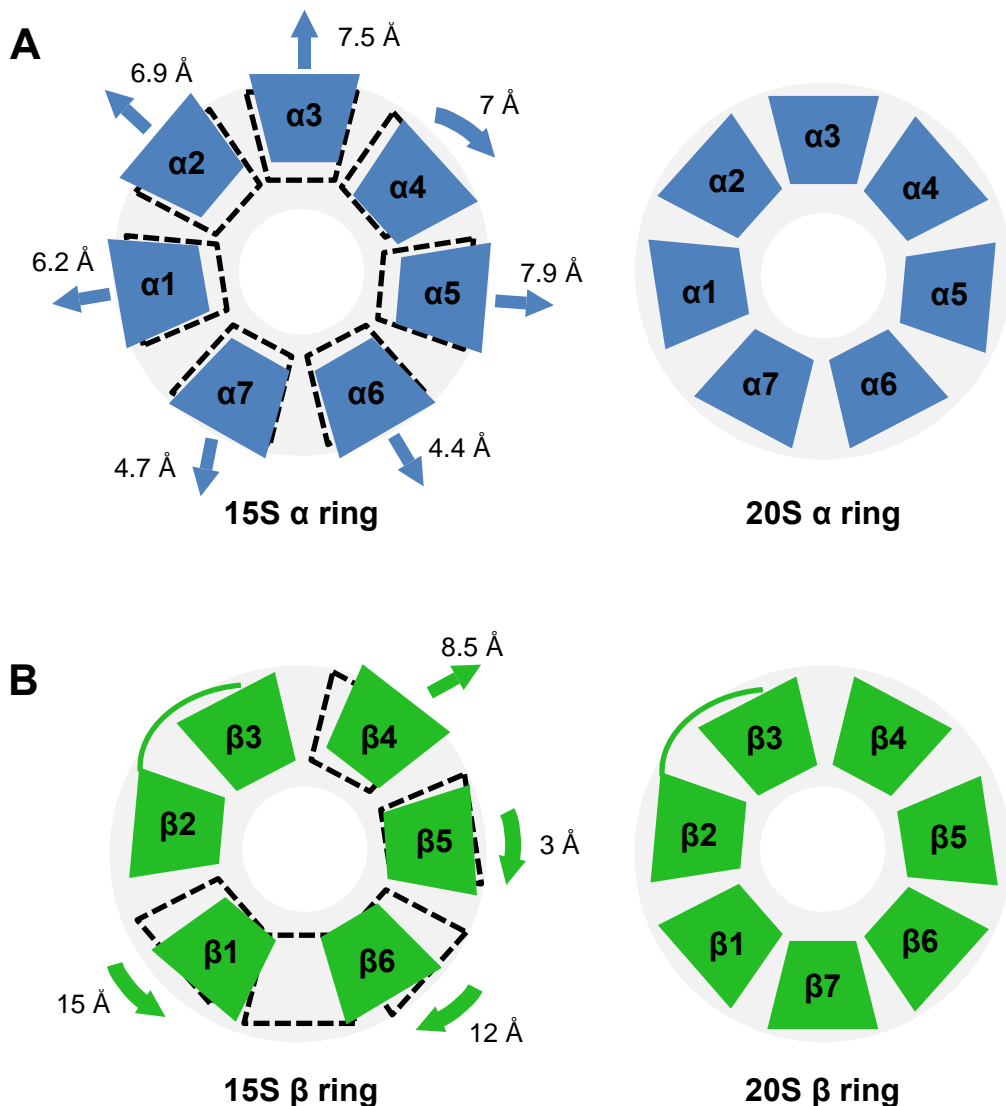


Figure 27: Schematic drawing of α (A) and β (B) subunit positions in the 15S complex compared to the positions in the mature 20S complex. Lateral subunit shifts are indicated by arrows and distances. The positions of the subunits in the mature complex are shown as dashed lines in the 15S complex for comparison.

Figure 28 provides an overview of cross-link distances according to the fit shown in figure 26. Only cross-links between neighboring subunits were taken into account for this analysis, as

there is no evidence of subunits swapping positions. Most of the cross-links are within a range of 15-24 Å. Several of the longer cross-links found in the range of 27-30 Å are formed between α subunits in the long loop region between H0 and H1, which is probably flexible and may increase the apparent cross-link distance. Only one cross-link was found that exceeds the 30 Å limit (between $\alpha 3$ K68 and $\alpha 4$ K39; 57.5 Å). However this cross-link was assumed to be formed between subunits of different complexes, as it also violates every other crystal structure fit that was explored, as well as the subunit arrangement in the mature complex.

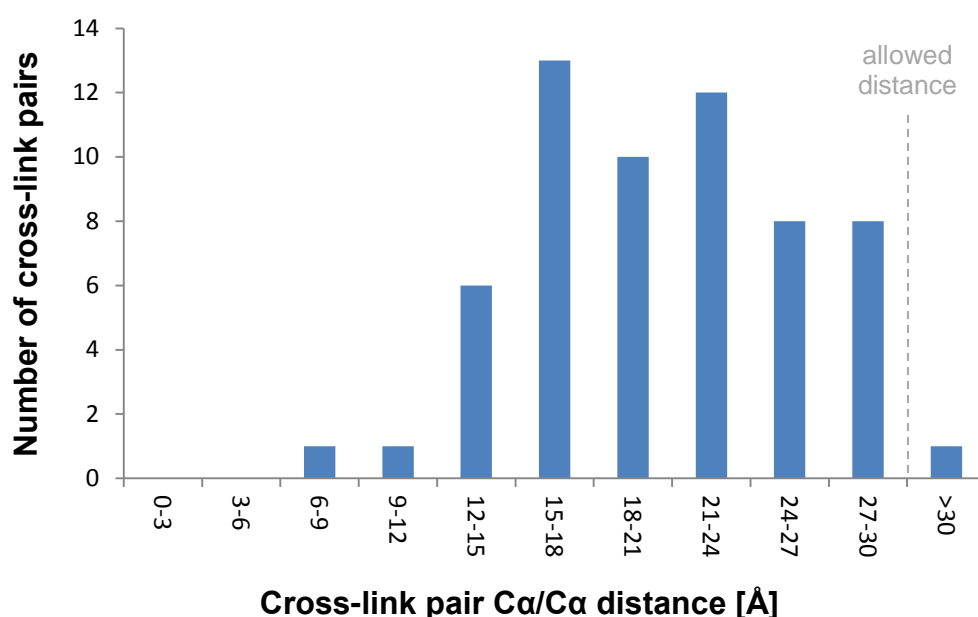


Figure 28: Chart showing the cross-link distances versus the number of cross-link pairs found in 15S and 15S^{GFP} samples investigated by MS. Only cross-links between neighbouring subunits were taken into account. Distances are measured between Ca atoms of cross-linked lysine residues according to the crystal structure fit shown in figure 26.

The fit of the 15S complex was also tested for the 15S^{GFP} EM map using the same coordinates for subunit crystal structures. In addition, the GFP crystal structure (Yang et al. 1996) was positioned in the EM density inside the 15S^{GFP} cavity. All GFP-related cross-links were drawn into the model to see if they can be accommodated. Figure 29 shows the resulting crystal structure fit. As expected from the eigenimage analysis and cross-linking data, GFP occupies a position inside the cavity, but only a few cross-links can be accommodated simultaneously by any possible fit of GFP, while the rest exceeds the 30 Å limit of lysine side chains cross-linked by DSS H12/D12. This points to GFP being flexibly attached, most likely near the subunits $\alpha 4$ and $\beta 4$, as this area harbors the Ump1 C-terminus to which GFP is fused. The flexibility is also observed in the electron density that can be attributed to GFP, as it is

too small to fit the entire molecule. The fit shown in figure 29 indicates that the cavity of the 15S^{GFP} complex is not big enough to allow large movements of the GFP molecule. Information from peptide mass fingerprinting (figure 8) and cross-linking analysis suggest that the subunits $\beta 5$ and $\beta 6$ are under-represented in the 15S^{GFP} sample. Missing subunits could widen the space for the GFP molecule to move. The incorporation of Ump1 into the complex is not affected by $\beta 5$ and $\beta 6$, as this event precedes the binding of those β subunits (Li et al. 2007). Nevertheless, missing subunits and increased flexibility of GFP do not account for a small number of cross-links that connect GFP to the outside of the 15S^{GFP} complex (as found in $\alpha 6$, $\beta 2$ and $\beta 4$). These cross-links can likely be explained by transient dimers that were observed in 15S samples (figure 15), but never specifically investigated in 15S^{GFP} samples.

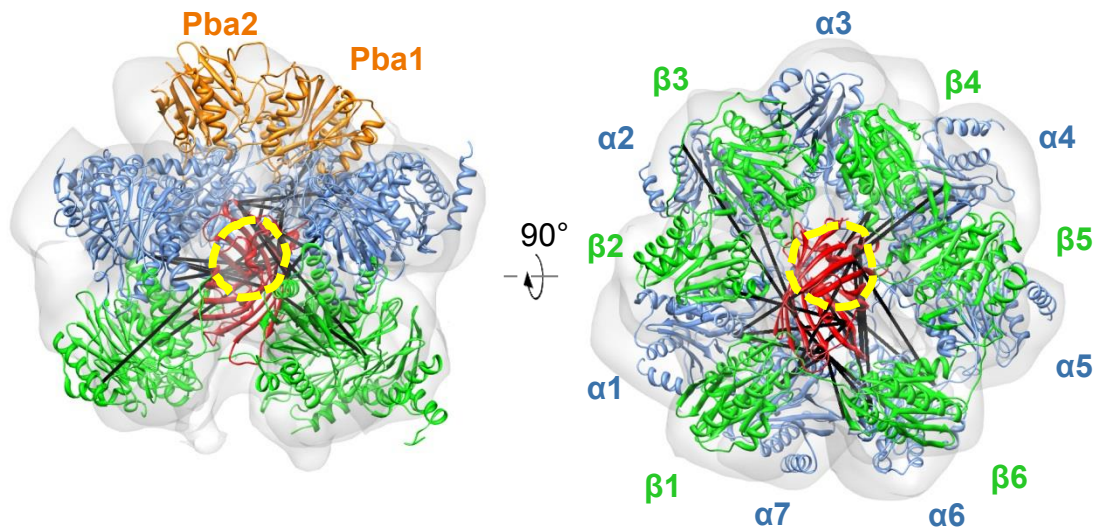


Figure 29: Fitting of crystal structure subunits into the 15S^{GFP} complex EM density. Apart from the additional GFP (red), the positions of crystal structure subunits are identical to the ones shown for the 15S complex in figure 26, with α subunits shown in blue, β subunits in green and Pba chaperones in orange. GFP-related cross-links are shown as black lines. The left image shows a cut open side view, in which $\alpha 3$ and $\beta 3$ have been deleted to improve visibility of the GFP molecule inside the cavity. The right image shows a bottom view of the complex, in which Pba1 and Pba2 have been deleted. Due to the transparency of the 15S^{GFP} EM map, the contours of the EM density corresponding to GFP have been additionally highlighted by a yellow dashed line.

6.5 Subunit fits for negative stain reconstructions - 20S *pre1-1* and 20S complexes:

The 20S *pre1-1* complex shows a structure that very much resembles that of the mature 20S, apart from the presence of extra density corresponding to the chaperones Pba1 and Pba2 (figure 30A). Rings are stacked in the typical $\alpha_7\text{-}\beta_7\text{-}\beta\text{-}\alpha_7$ fashion, with all subunits present in the complex. Comparing the height of the mutated 20S *pre1-1* complex with that of the wild type 20S reveals that the late stage precursor is 8.5 Å longer than the mature complex (figure 30C). Core densities of the 20S *pre1-1* complex are shifted away from the equatorial plane, although that shift is much more pronounced in the α subunits than in the β subunits. Additionally, there is a larger degree of separation between β subunit EM densities of both rings in the mature 20S (figure 30B), while the same subunits are more connected by EM density in the 20S *pre1-1* precursor complex. Due to the apparent outward shift of the two halves of the proteasome, the 20S *pre1-1* EM map has been fitted with half-proteasomes derived from the reconstituted 20S-Pba1-Pba2 complex (Stadtmueller et al. 2012), while the mature 20S could be fitted with the crystal structure of the complete *S. cerevisiae* 20S proteasome (Groll et al. 2000). The α ring of the 20S *pre1-1* complex resembles that of the mature 20S, as indicated by identical subunit fits for the mutant and wild type complexes. Just like in the wild type 20S complex, the pore between the α subunits is closed in the 20S *pre1-1* complex. In contrast to the disordered arrangement of the β subunits in the 15S complex, the same subunits appear more ordered in the 20S *pre1-1* complex. While their overall arrangement looks very similar to the one found in mature 20S, the subunits $\beta_4\text{-}\beta_6$ are slightly shifted outwards, away from the pore axis. Combined with the 8.5 Å gap at the β/β interface, this gives the middle of the 20S *pre1-1* complex a bulkier and less defined appearance than the wild type 20S. One end of the 20S *pre1-1* complex is occupied by a Pba1-Pba2 heterodimer that shows a different arrangement than in the 15S precursor complex. Details are described in section 6.6.

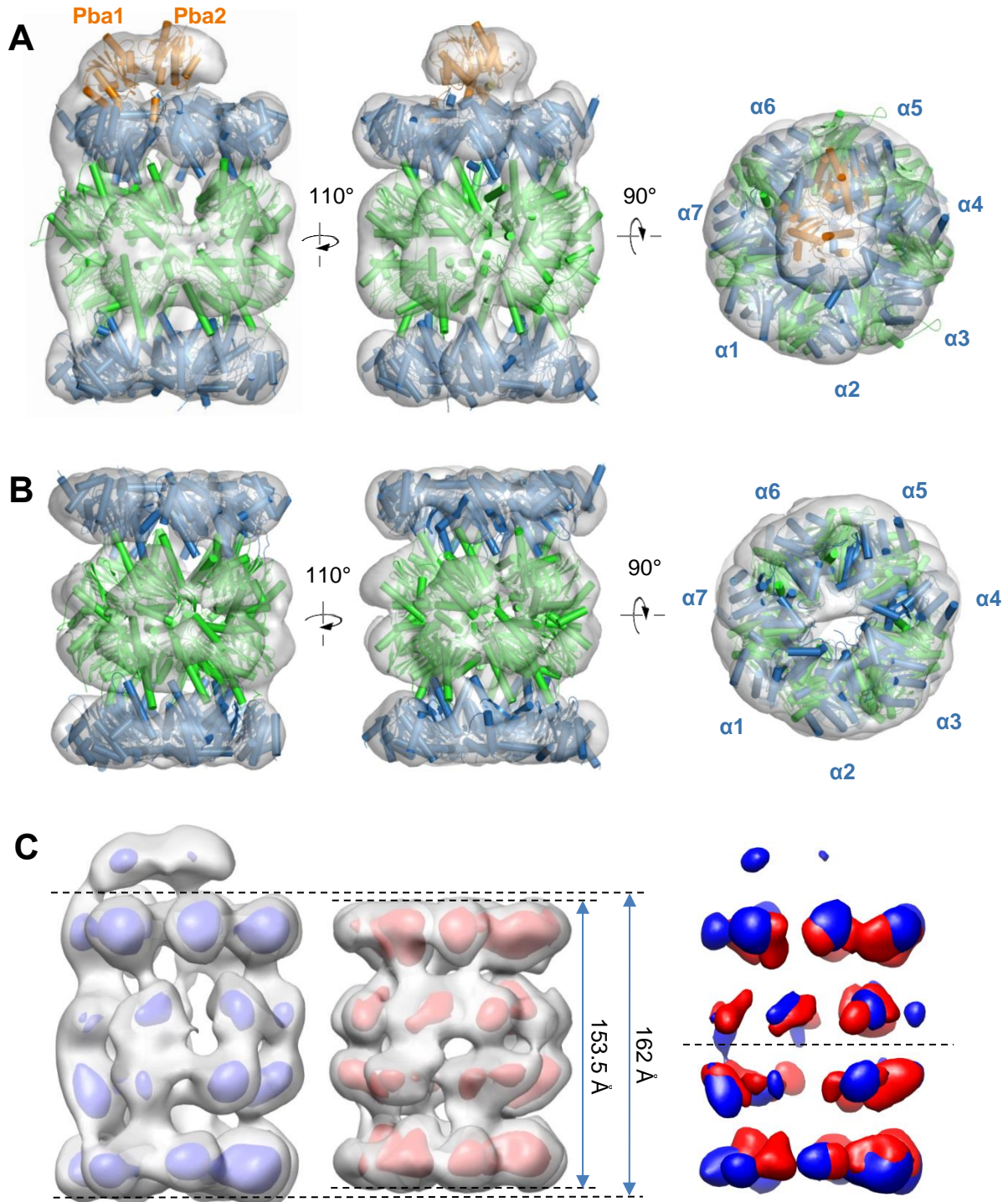


Figure 30: Crystal structure fits and comparisons between mutated 20S *pre1-1* and wild type 20S complexes. **(A)** The 20S *pre1-1* complex EM map has been fitted with two half proteasomes independently derived from the crystal structure of the reconstituted 20S-Pba1-Pba2 complex (Stadtmueller et al. 2012). Individual subunit fits were performed for β subunits to better accommodate the EM density. α subunits are shown in blue, β subunits in green and Pba chaperones in orange. **(B)** The 20S EM map has been fitted with the crystal structure of the *S. cerevisiae* 20S proteasome (Groll et al. 2000). No individual subunit fits were performed. Subunits are coloured same as in (A). **(C)** Size comparison of the 20S *pre1-1* and 20S complexes rendered at the threshold corresponding to their respective molecular weight. Core densities of 20S *pre1-1* are shown in blue, while those of the 20S complex are shown in red. The equatorial plane has been indicated by a dashed line in the rightmost image.

6.6 Localization of the Pba1-Pba2 heterodimer in the 15S and 20S *pre1-1* precursor complexes:

As seen in the previously published crystal structure of *S. cerevisiae* 20S reconstituted with Pba1-Pba2 (Stadtmueller et al. 2012), the heterodimer was localized in the opening of the α ring. The mammalian orthologues of Pba1 and Pba2 were reported to only work in tandem, hence it is assumed that this configuration also exists in the 15S precursor complex. Apart from the fact that two distinct densities were found in the EM map, this assumption is also backed up by a cross-link found between Pba1 K213 and Pba2 K65 at the interface between the two proteins (figure 31A). Fitting crystal structure subunits into the EM density reveals a much closer association of Pba1-Pba2 with the rest of the precursor complex than seen in the reconstituted complex that led to the crystal structure. Continuous density exists between all of the α subunits and Pba1-Pba2, except for $\alpha 1$. The strongest connections are observed near the subunits $\alpha 3$ and $\alpha 4$, as well as $\alpha 6$ and $\alpha 7$. While the low resolution of the EM map does not allow for identification of specific interactions at atomic level, most of them seem to be mediated by the N-terminal regions of the α subunits. However, cross-links between Pba1-Pba2 and α subunit N-termini were not found in any of the complexes examined by cross-linking analysis, probably because these protein regions are almost completely devoid of lysine residues. Notably, the special arrangement of Pba1-Pba2 in relation to the α subunit N-termini leaves the α ring pore open beneath the chaperones, while the latter effectively “plug” the pore. This can be seen in the cut open side view of the 15S complex in figure 32. The main interaction surface between the heterodimer and the α ring is between the N-terminal regions of $\alpha 6$ and $\alpha 7$, where the HbYX motif of Pba2 binds. Not only does this site exhibit strong electron density in the 3D reconstruction, it also displays several cross-links between $\alpha 7$ (K57 and K167) and the Pba2 C-terminus (K261) (figure 31B). Whether $\alpha 6$ takes part in binding the Pba2 C-terminal HbYX motif is uncertain. On the one hand the EM density allows fitting $\alpha 6$ in a position that enables binding, on the other hand no cross-links have been observed between Pba2 and $\alpha 6$ (although Pba2 cross-links to the equally close $\alpha 7$). Curiously, no interaction between the Pba1 HbYX motif and the $\alpha 6$ subunit is observed in the 15S model, neither in EM density nor through cross-links. This is surprising since $\alpha 6$ has been reported to be a part of a canonical HbYX binding pocket, forming a hydrogen bond between the pocket lysine K62 and the C-terminal carboxyl group of the HbYX motif (Stadtmueller et al. 2012). In fact the fit of Pba1 and $\alpha 6$ in the EM density places $\alpha 6$ 9.5 Å away from its location found in the crystal structure of the reconstituted 20S-Pba1-Pba2 complex (figure 31B). Instead the HbYX motif of Pba1 seems to solemnly interact with $\alpha 5$, which can form

hydrogen bonds between L21 and E25 and the HbYX motif tyrosine side chain (Stadtmueller et al. 2012). Pba1 also interacts with the α ring in locations other than the HbYX motif binding sites. A cross-link was found between Pba1 K91 and $\alpha 3$ K51, indicating that this α subunit is close to the chaperone (figure 31C). The exact contact points remain to be investigated, as the corresponding residues of Pba1 are not resolved in the crystal structure. Electron density connecting Pba1 with $\alpha 3$ and $\alpha 4$ is observed in the EM map of the 15S complex, although it is unclear whether it corresponds to the part missing from the Pba1 crystal structure.

In contrast to the assembly of the 15S complex, the Pba1-Pba2 heterodimer of the 20S *pre1-1* precursor complex shows an arrangement that resembles the one found in the crystal structure of the reconstituted 20S-Pba1-Pba2 complex. This is in large part due to a tighter α ring arrangement, which changes the overall geometry and consequently influences binding modes for Pba1-Pba2. The chaperones have moved out of the pore by ~ 10 Å and are now found on top of the α ring (figure 32). In fact the crystal structure subunits of Pba1-Pba2 and the α ring can be fitted into the electron density without major readjustments. Not only does the heterodimer no longer occupy the pore, it is also moved in the direction of $\alpha 5$ by ~ 16 Å, which gives the α ring and its chaperones a more asymmetric appearance compared to the 15S complex, where the chaperones were located almost in the middle of the ring. In contrast to the 15S complex, the pore in the 20S *pre1-1* complex is closed without involvement of the chaperones. Instead the tighter arrangement of the α subunits allows their N-termini to close the pore, just like in the mature 20S complex. Concerning the interaction of the α subunits and the HbYX motifs of the chaperones, continuous electron density is found between Pba1 and $\alpha 5$ near the HbYX motif binding site between subunit $\alpha 5$ and $\alpha 6$. Density connecting Pba2 to the subunit $\alpha 7$ as it was found in the 15S complex is not observed in the 20S *pre1-1* complex unless the threshold used to render the map is lowered. In contrast, the apparent interaction of Pba1 and $\alpha 3$ that was observed through cross-linking data in the 15S complex is likely to be preserved in the 20S *pre1-1* complex, as EM density is extending from the chaperones to the $\alpha 3$ subunit (figure 32).

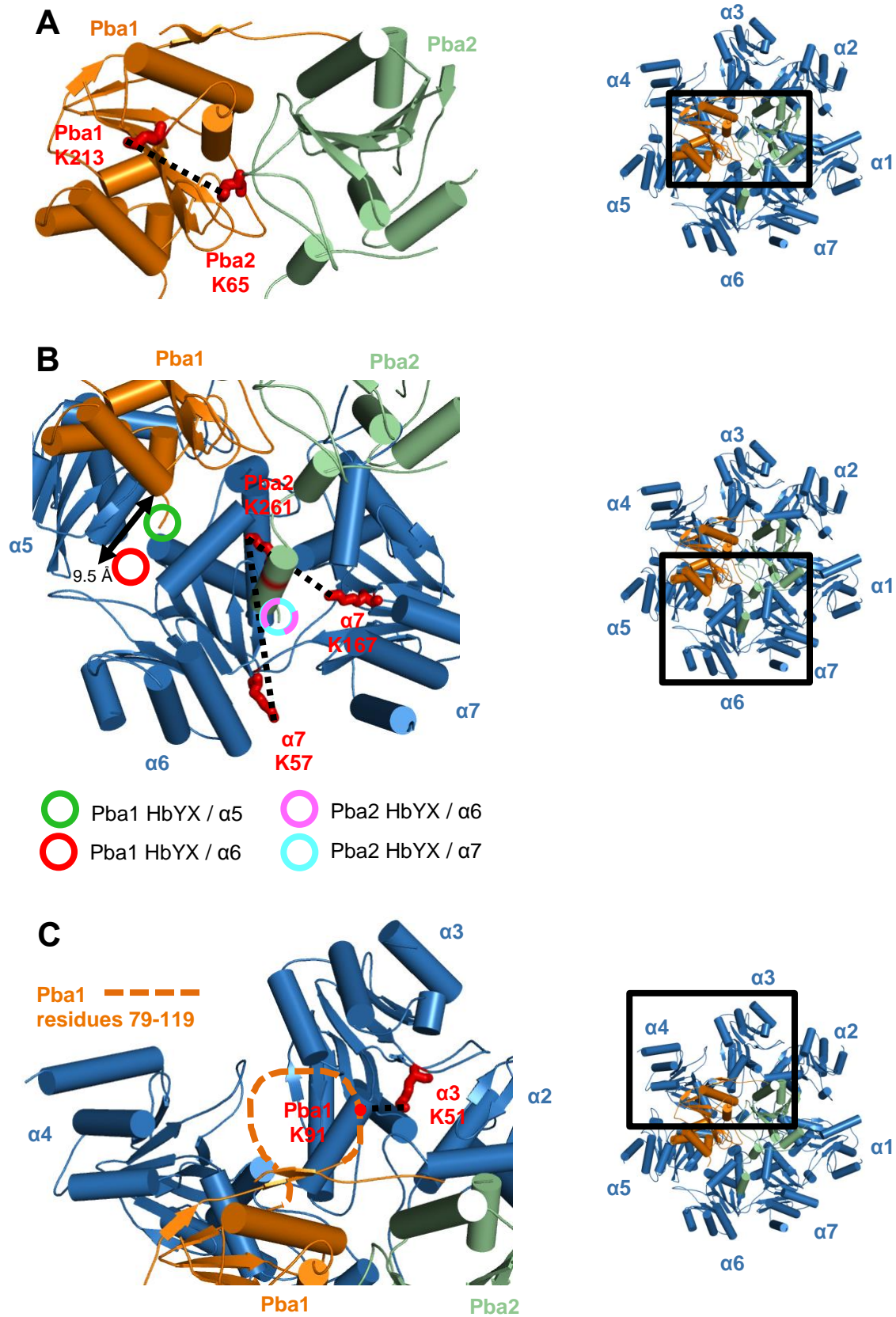


Figure 31: Cross-links between chaperones and α subunits of the 15S complex. The rightmost images indicate which part of the complex is displayed in detail on the left. Cross-links are shown as black dashed lines, while cross-linking lysine residues are displayed as red sticks (A) The cross-link at the interface between Pba1 and Pba2 shows that it acts as a heterodimer in the 15S complex. (continued)

Figure 31 (continued): (B) Cross-links between $\alpha 7$ and the Pba2 C-terminus. The positions of Pba1 and Pba2 HbYX motifs relative to the subunits $\alpha 5$, $\alpha 6$ or $\alpha 7$ (as found in the reconstituted 20S-Pba1-Pba2 complex crystal structure) are indicated by coloured circles. (C) Cross-link of $\alpha 3$ to the unstructured part of Pba1. The missing amino acids of Pba1 are indicated as an orange dashed line.

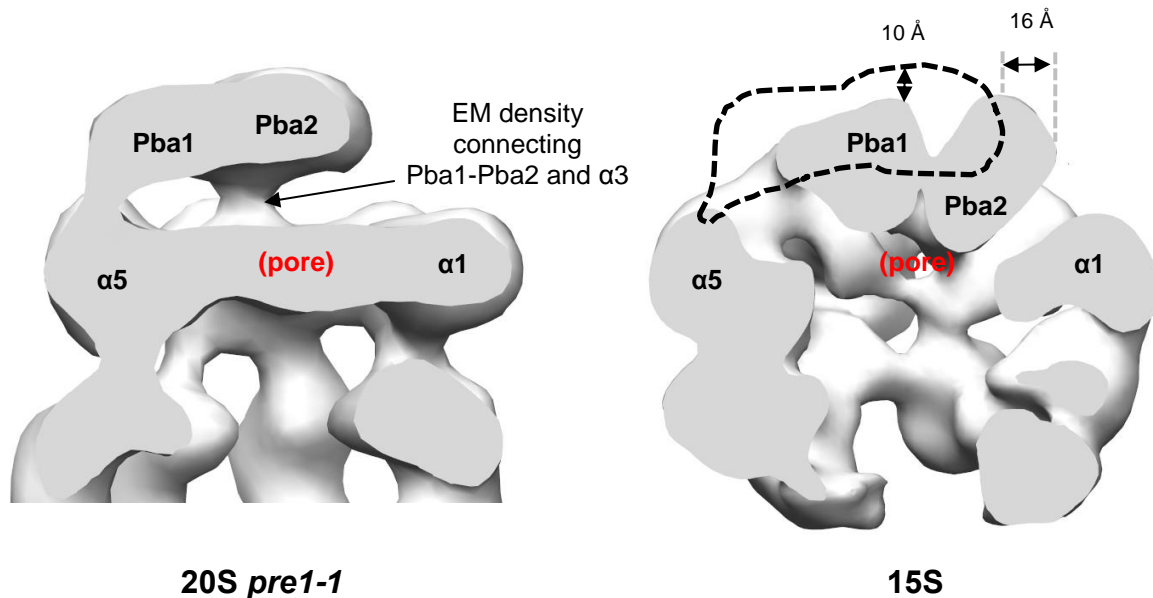


Figure 32: Cut open side views of the 20 *pre1-1* and 15S complex showing differences in the arrangement of the Pba1-Pba2 densities. The black dashed line in the 15S reconstruction indicates the position of the Pba1-Pba2 densities as found in the 20S *pre1-1* complex.

6.7 Localization of Ump1 in the 15S precursor complex:

Using cross-linking and MS analysis of the 15S and 15S^{GFP} complexes the chaperone Ump1 was localized in the precursor complex. Figure 33 shows the projected trajectory of the Ump1 amino acid chain. Starting at the C-terminus of Ump1, the lysine residues K88 of $\alpha 4$ and K90 of $\beta 4$ both cross-link to K144 of Ump1, which is only 4 residues away from its C-terminal end. These cross-links were found both in the complex with and without the GFP fusion protein, indicating that the addition of the GFP moiety does not hinder regular incorporation of Ump1. The location of the Ump1 C-terminus near $\alpha 4$ and $\beta 4$ was confirmed by positioning GFP in the EM map of 15S^{GFP} in accordance with cross-linking data (figure 29). In this position GFP is located in the middle of the complex, pointing its N-terminal end toward the location of the Ump1 K144 cross-links. The lysine in position 141 of Ump1 does not form inter- or intra- cross-links but is restricted to mono-links. Further inter-protein cross-links were found in the $\alpha 1$ lysines K107 and K98, connecting to Ump1 K113 and K83, respectively. This interaction was verified by binding studies of isolated proteins *in-vitro*,

which found a strong binding between $\alpha 1$ and Ump1 (Dohmen lab, unpublished data). While the residues of $\alpha 1$ that interact with Ump1 are located close to each other in the same helix (H1), the interacting amino acids of Ump1 are 30 residues apart, which leaves room for binding to neighbouring subunits. Moving further up the amino acid chain, K58 of Ump1 was found to cross-link both to the C-terminal K144 mentioned earlier and K156 of GFP (not shown in figure 33). The latter is located near the N-terminal end of GFP, so it is in close proximity to the C-terminus of Ump1, where aforementioned cross-links were found. The most N-terminal lysine found to cross-link to regular proteasome subunits was K19 of Ump1, which links to K91 of $\beta 6$. This puts the N-terminal end of Ump1 near the opening of the β -ring pore. Cross-links for Ump1 K10 were not observed, but it is entirely possible that this end of Ump1 is pointing out of the pore and into the lumen. Evidence for this orientation of the N-terminus comes from the purification of the 15S complex via a FLAG-his₆ tag on the N-terminal end of Ump1. This end has to be highly accessible if the 15S complex is to be purified intact. To confirm the accessibility, the N-terminus of Ump1 was labelled with 5 nm Nanogold particles (see figure 34) covalently linked to Ni²⁺ NTA. While Ump1 cannot be directly localized using this method, the gold particles nevertheless bind complexes and seem to force them into a side view orientation on the grid, which is only possible if the Ump1 N-terminus is exposed at the opening of the β ring pore. With such an exposed N-terminus a mono cross-link was expected for K10 of Ump1 but none was found in the data. Similarly, although it has been described earlier in the literature for the human ortholog hUmp1 (Heink et al. 2005), an interaction between yeast Ump1 and $\beta 5$ was not observed in the cross-linking data.

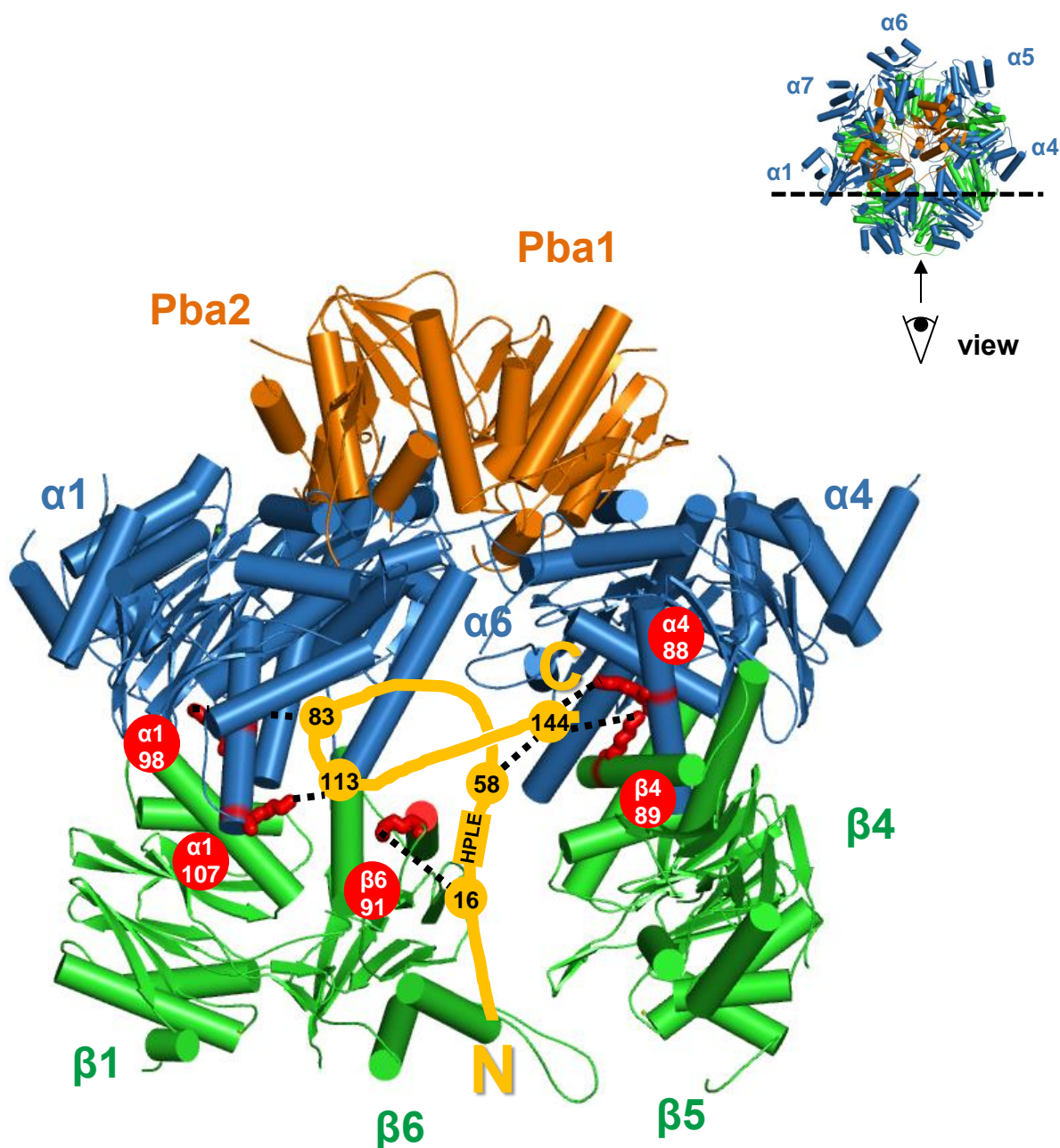


Figure 33: Predicted trajectory of the Ump1 amino acid chain inside the 15S complex. The 15S complex is displayed as a cut open side view (as indicated on the upper right) to better show the Ump1 trajectory on the inside of the complex. α 2-3 and β 2-3 have been deleted from this image. α subunits are coloured in blue, β subunits in green and Pba chaperones in orange. The subunit arrangement is analogous to the one shown in figure 26. The Ump1 amino acid chain is shown as a yellow line (not to scale) with the N- and C-terminus indicated by letters. Encircled numbers within the chain indicate amino acid positions where cross-links have been found. Cross-links are shown as dashed black lines. The position of the conserved HPLE motif in the amino acid chain of Ump1 is also shown. The cross-linking residues from α and β subunits are displayed as red sticks with the subunit and amino acid residue number shown in red circles next to them.

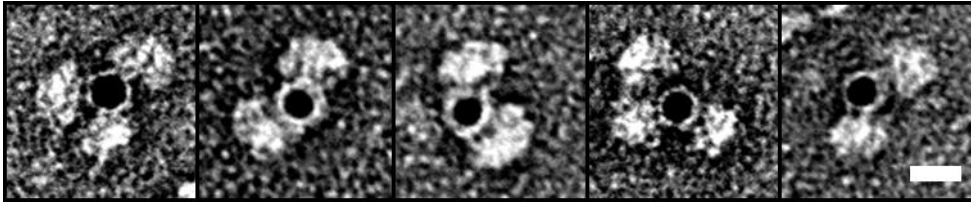


Figure 34: Examples of 15S complexes bound to 5 nm Nanogold which is covalently linked to Ni^{2+} NTA. Protein complexes were used at a concentration of 50 $\mu\text{g}/\text{ml}$ and mixed with 10 nmol/ml Nanogold particles at a ratio of 50:1. Nanogold particles are visible on the grid as thick black dots. The linker distance between the Nanogold particles and the FLAG- his_6 tag on the N-terminus of Ump1 is assumed to be less than 1.5 nm (according to the manufacturer). The bar corresponds to a distance of 10 nm.

7. Discussion:

7.1 The role of Pba1-Pba2 in 20S biogenesis:

7.1.1 Alterations of the HbYX motif binding pockets influence Pba1-Pba2 binding modes:

Before discussing the role of Pba1-Pba2 in 20S biogenesis, it is advisable to take a look at the binding sites that these chaperones occupy in the early and late precursor complexes and how the binding modes might change. In the 15S complex, all α subunits are moved farther away from the centre of the α ring compared to the mature 20S. This indicates that canonical features of binding pockets for HbYX motifs between the subunits have been altered or abolished. Due to the larger diameter of the ring, proteins that interact with more than one pocket (like the Pba1-Pba2 heterodimer) would have to adapt to the increased distance and binding conditions. That's why it comes as no surprise that proteins binding to these pockets may adopt different binding modes and arrangements than in the mature complex. In the crystal structure of the reconstituted 20S-Pba1-Pba2 complex, Pba1-Pba2 are far away from the α ring, only showing interactions with $\alpha 5$, $\alpha 6$ and $\alpha 7$ through their HbYX motifs (Stadtmueller et al. 2012). The other subunits of the α ring are not contacted directly in that structure. In the 15S precursor complex the C-terminal helix of Pba2 is bound by $\alpha 7$, as indicated by cross-links between these two subunits near the HbYX binding site and the strong EM density observed in this location. Thus an interaction of the Pba2 HbYX motif with the pocket lysine ($\alpha 7$ K65) as seen in the reconstituted complex is possible. As described above the binding of the Pba1 HbYX motif is different in the 15S complex compared to later stages of maturation. While the EM density suggests that the C-terminus of Pba1 is located close to $\alpha 5$, no interaction can be observed with $\alpha 6$. The latter is shifted toward $\alpha 7$ and away from $\alpha 5$, which increases the distance between $\alpha 5$ and $\alpha 6$ to a degree that only one of them can contribute to binding Pba1. This assumption is backed up by cross-linking data, which shows multiple cross-links connecting neighbouring α subunits throughout the ring via flexible loops found between helices H0 and H1. No such connection can be observed for $\alpha 5$ and $\alpha 6$, indicating that the distance between them is increased, just as it was observed in the EM map. Additional evidence about the binding of the chaperone heterodimer to specific α subunits comes from the human orthologues PAC1-PAC2, which could only interact with $\alpha 5$ and $\alpha 7$ but not with $\alpha 6$ *in-vitro* (Hirano et al. 2005).

In the 20S *pre1-1* complex the HbYX binding sites are fully restored due to a tighter arrangement of α subunits that resembles the one found in the mature 20S. In addition to the

canonical lysine residues already mentioned above, the residues lining these pockets can also contribute hydrogen bonds for binding HbYX motifs (Stadtmueller et al. 2012). In this state the binding of the chaperones seems to be mainly mediated by Pba1, as shown by mutational analysis of HbYX motifs in Pba1 and Pba2. This is in line with observations made in the EM map of the 20S *pre1-1* complex, which shows a strong connection between Pba1 and $\alpha 5$.

Summarizing the interpretation of data from the 15S and 20S *pre1-1* complexes, the HbYX motif binding sites are altered in the early precursor due to the increased distances between individual α subunits. Binding of the chaperone HbYX motifs seems to be mediated by $\alpha 5$ and $\alpha 7$ alone, while $\alpha 6$ takes no part in the binding. In the late precursor the HbYX motif binding sites are restored and $\alpha 6$ can bind the chaperones.

7.1.2 The position of Pba1-Pba2 in the precursor complexes changes during maturation:

One of the most striking features of Pba1-Pba2 in the 15S precursor complex is the way it is arranged in the α ring pore. It is located in the middle of the pore, contacting all but one α subunit directly. In contrast, none of the available crystal structures of reconstituted 20S-Pba1-Pba2 (Stadtmueller et al. 2012) and $\alpha 5$ -Pba3-Pba4 complexes (Yashiroda et al. 2008) explain how the subunits $\alpha 1$, $\alpha 2$ and $\alpha 3$ are contacted in the proteasome precursor complex. In the case of $\alpha 3$ this is a particularly important question, as it is the only non-essential subunit of the 20S proteasome and has to ensure its incorporation under non-stress conditions. In contrast, the 15S EM map shows contacts of Pba1-Pba2 to $\alpha 2$ and $\alpha 3$. The proximity of Pba1 to $\alpha 3$ could be independently verified by cross-linking analysis. While no direct contact with $\alpha 1$ is observed in the EM map, the chaperone heterodimer is nevertheless in close proximity of $\alpha 1$ and a lot closer than in the reconstituted 20S-Pba1-Pba2 complex (see figure 32).

Due to the increased diameter of the α ring and their central location, Pba1-Pba2 likely contact the N-termini of α subunits. Unfortunately most of these contacts could not be independently validated by cross-links due to the lack of lysines in the α subunit N-termini (with the aforementioned exception of the Pba1/ $\alpha 3$ interaction slightly downstream of the $\alpha 3$ N-terminus). Possible contacts to the α subunit N-termini are in line with previously published data from the archaeal system showing that the interaction between Pba chaperone orthologues and the α ring is abolished, once the first 13 amino acids are deleted from α subunits (Kusmierczyk et al. 2011). It also clearly argues against the claim that the elevated Pba1-Pba2 dimer as seen in the late precursor 20S *pre1-1* or reconstituted 20S-Pba1-Pba2 complexes can help to assemble the α ring, because the α subunit N-termini seem to play little to no role in those structures with respect to heterodimer binding. Instead the latter represents

an artificial reconstruction of the last stage in maturation. This assertion is backed up by the EM reconstruction of the 20S *pre1-1* complex, which bears a resemblance to the α subunits and Pba chaperones of the reconstituted complex.

While the physical presence of Pba1-Pba2 on top of the pore is somewhat mitigated in the 15S complex, it would still be able to inhibit interaction with unwanted activators trying to interact with the α ring by blocking the α ring surface and certain HbYX motif binding sites. The dimerization of α rings that has been reported in strains lacking Pba1-Pba2 would also be blocked by the presence of the chaperones.

In summary, it appears as though the Pba1-Pba2 chaperones possess different binding modes for different stages of 20S biogenesis. In the beginning of α ring assembly, they occupy the middle of the α ring pore, binding these subunits mostly through interactions with their N-termini, while the binding mediated by HbYX motifs could be confirmed for Pba2 but seems somewhat mitigated in Pba1. In later stages of 20S biogenesis the α ring contracts, restoring binding pockets for HbYX motifs to full function. At this stage the chaperone heterodimer is pushed out of the pore and interactions with α subunit N-termini are mostly abolished (with the notable exception of the Pba1/ α 3 interaction, which is still observed in the EM map of the 20S *pre1-1* complex). Instead Pba1-Pba2 can now rely on the interaction of its HbYX motifs with the binding pockets between α 5/ α 6 and α 6/ α 7 to keep it bound to the nascent 20S. Just like in the 15S complex, the Pba1-Pba2 presence in the late precursor can block interactions of the α ring with unwanted activators, until the 20S has completed its maturation.

7.1.3 Possible cross-talk between the α and β ring:

The 3D reconstructions of the early precursor 15S and the late precursor 20S *pre1-1* suggest a cross-talk between the α and β ring of the nascent proteasome. In the 15S complex the α ring diameter is increased compared to the late precursor and the mature complex, even though there is no obvious maturation event happening in the α subunits afterwards (e.g. like pro-peptide cleavage as observed in certain β subunits). Instead it is possible that the tightening of the α ring is a result of events in the β ring, like the β subunits adopting a more ordered conformation for dimerization of two half-proteasomes. The latter is suggested by a transformation of the β ring from a relatively disordered state in the 15S complex to a more ordered one in the late precursor 20S *pre1-1*, which is accompanied by the aforementioned tightening of the α ring. A cross-talk between the two rings has been reported before, when it was observed that the binding of the activator Blm10 to the α ring can have effects on the rate of β subunit processing of pro-peptides, even though they never contact each other directly

(Fehlker et al. 2003). Details on how this cross-talk is accomplished are missing, although something similar has been described for the catalytic β subunits and the 19S regulatory particle, where signals are presumably passed through the α subunits (Kleijnen et al. 2007).

The 20S *pre1-1* complex delivers another example of a possible cross-talk between α and β subunits. Micrographs displaying particles from the 20S *pre1-1* sample show a high degree of heterogeneity (see figure 16C). Uncapped, single-capped and double-capped 20S *pre1-1* complexes are present in the sample. Western blots of the 20S *pre1-1* sample using antibodies directed against β 2 and β 5 show a mixture of mature and un-matured β subunits (Dohmen lab, unpublished data). Using previously published data from archaea it can be theorized that the binding of Pba1-Pba2 to the proteasome is directly related to the state of the catalytic β subunits active sites. The Pba orthologues in archaea preferentially bind to precursor assembly intermediates of 20S (Kusmierczyk et al. 2011). The same was observed for mature 20S treated with Z-Leu₃-vinyl sulfone or *clasto*-lactacystin, which both inhibit catalytic β subunits. In contrast, untreated mature 20S was unable to bind the chaperones. It appears as though the catalytic β subunits more readily bind Pba chaperones when pro-peptides or inhibitors are present in the active sites. In a related matter, the reconstituted complex of 20S-Pba1-Pba2 from *S. cerevisiae* was unable to crystallize without adding the inhibitor MG132 (Stadtmueller et al. 2012). This points to the catalytic β subunits having an influence on Pba1-Pba2 binding, although it is unclear how this is accomplished because the 20S from the reconstituted 20S-Pba1-Pba2 crystal structure only shows modest structural changes compared to unbound 20S. A possible preference of Pba1-Pba2 for precursor complexes means that mature 20S complexes can exist in the 20S *pre1-1* sample, but they would have already shed their Pba1-Pba2 chaperones. This would in turn lead to the high degree of heterogeneity observed in EM micrographs and the low occupancy of Pba1-Pba2.

7.2 The role of Ump1 in 20S biogenesis:

7.2.1 The predicted structure of Ump1:

The trajectory map (figure 33) shows that Ump1 crosses the pore twice, contacting subunits on both sides and in both the α and β ring. Contacts to α subunits have previously only been observed in its human ortholog hUmp1 but were unheard of in yeast, where it was assumed that Ump1 exclusively coordinates the incorporation of β subunits, as it follows β 2, β 3 and β 4 into the ring. The stretched out conformation of Ump1 and the fact that there is no electron density visible along the projected path fits the description of an intrinsically unstructured

protein with little secondary structure elements, as predicted in two recent studies that were unable to determine secondary or tertiary structure elements of isolated Ump1 using a wide array of biophysical techniques (Sá-Moura et al. 2013; Uekusa et al. 2013). It also explains how Ump1 can be degraded by the nascent proteasome without dissolving its tertiary structure first. Similar examples of the 20S proteasome being able to degrade naturally unfolded polypeptides have been found in α -synuclein, tau and p21^{Cip1} (Tofaris et al. 2001; David et al. 2002; Sheaff et al. 2000). Localizing Ump1 in the pore fits previous biochemical studies in which Ump1 was inaccessible to polyclonal antibodies as soon as two proteolytically inactive half-proteasomes dimerized (Ramos et al. 1998).

The cross-linking data does not provide evidence that Ump1 dimerizes when bound to the precursor complex (as observed for unbound Ump1 in Sá-Moura et al. 2013). In addition, SDS-PAGE analysis does not suggest an enrichment of the protein compared to other 15S subunits. Ump1 was proposed to dimerize via a disulphide bond between two C115 residues, which according to cross-linking analysis are located deep inside the pore close to the α 1 subunit. The Ump1-GFP-HA fusion protein incorporated into 15S^{GFP} complexes would not be able to dimerize as fitting two GFP moieties into the internal cavity of the complex is impossible due to steric clashes.

7.2.2 Correlating structure and function of parts of Ump1:

When Ump1 is incorporated into 13-16S precursor complexes, it was found to be protected against trypsin digestion with the exception of a ~5 kDa N-terminal part (Ramos et al. 1998). Coincidentally, this part of Ump1 seems to be dispensable with respect to the proteins incorporation into the complex, as the first 50 amino acids can be substituted without influencing binding to the nascent proteasome in mammalia, although they are still needed for forming fully functional 20S (Burri et al. 2000). Purifications of the 15S^{GFP} complex yielded a similar result. Here, an N-terminal fragment of comparable size is cleaved off of Ump1 in a small subset of proteins, presumably by cellular proteases (see figure 8). Even without these amino acids, Ump1 is still incorporated into the 15S complex as evidenced by analytical ultracentrifugation (see figure 9).

Position 51-54 of Ump1 houses the HPLE motif. While this motif is conserved in yeast and some mammals, its deletion did not affect hUmp1 binding in human cells, so it is clearly not responsible for binding the chaperone to the nascent proteasome (Sá-Moura et al. 2013; Burri et al. 2000). The residues downstream of the HPLE motif were found to be the more conserved between yeast and mammals, while the first 50 amino acids are not only more

poorly conserved but also predicted to be more flexible than the rest of the molecule (Sá-Moura et al. 2013). This prediction is in line with findings from cross-linking analysis and Nanogold labelling, which clearly showed the Ump1 N-terminus to exit the β ring pore, possibly pointing into the lumen. The observation that these residues are dispensable for Ump1 incorporation, yet necessary for the formation of fully functional 20S, lead to the hypothesis that they are involved in the dimerization of half-proteasomes. Indeed, attaching FLAG-his₆ or GFP moieties to the N-terminus of Ump1 appears to stall proteasome biogenesis and lead to an accumulation of 15S precursor complexes, eventually killing the cells (Dohmen lab, unpublished data).

Since the N-terminus of Ump1 is apparently not involved in binding of proteasome precursors, this places special emphasis on the parts of Ump1 that are found deeper inside the pore. These residues (namely 51-148) interact with the subunits $\alpha 1$, $\alpha 4$, and $\beta 4$. They cross the pore twice, as evidenced by the cross-link of Ump1 K58 that connects back to its C-terminus (K144). Due to this “loop-like” trajectory, Ump1 may feature many more contacts with α and β subunits along the wall of the 15S complex cavity that could not be detected in cross-linking analysis due to the lack of cross-linkable lysines in the right locations. In the mammalian ortholog hUmp1 it was found that deleting the residues 67-90 renders the protein unable to bind to the proteasome (Burri et al. 2000). Deletion of residues 68-72 also severely affected hUmp1 binding. According to cross-linking analysis of yeast Ump1, these residues belong to the part of Ump1 that interacts with $\alpha 1$ or are in the close vicinity of the neighbouring subunits $\alpha 6$ and $\alpha 7$. The subunit $\alpha 1$ was also found to be a strong binder of Ump1 when using isolated proteins *in-vitro* (Dohmen lab, unpublished data), so the interaction between these proteins seems to be of great importance for Ump1 binding. However it is likely that yeast Ump1 needs secondary binding partners, as the protein is only incorporated into the precursor complex if $\beta 4$ is present in addition to $\alpha 1$. The interaction between $\beta 4$ and the Ump1 C-terminus was observed in all cross-linked samples investigated by MS and may serve as an explanation why the binding to $\alpha 1$ alone does not lead to an earlier incorporation into the precursor complex. The C-terminus of the chaperone was found to be dispensable in hUmp1 (Burri et al. 2000), but mammalian orthologues bind to precursor complexes much earlier than in yeast and is not dependent on $\beta 4$, so these findings are not applicable to yeast Ump1.

While the cross-linking data does not provide direct evidence of yeast Ump1 interacting with the pro-peptide of $\beta 5$, this interaction has been previously shown for its human ortholog hUmp1 (Heink et al. 2005). Interestingly, cross-links between the $\beta 5$ pro-peptide and the

subunit $\alpha 6$ have been found on the inside of the 15S complex cavity where the C-terminal part of Ump1 resides (see figure 33 for the trajectory map). The trajectory map shows that the part of Ump1 closest to $\alpha 6$ is also the one found to be important for the hUmp1/ $\beta 5$ pro interaction in human cells. This is intriguing as it suggests a connection between Ump1 and $\beta 5$ pro inside the cavity. Li et al. (2007) have speculated that $\beta 5$ pro is the principle driving force behind the dimerization of half-proteasomes, but Ump1 keeps it in a conformation that prohibits dimerization until the last subunit $\beta 7$ is incorporated into the half-proteasomes. The cross-linking data of Ump1 and $\beta 5$ pro seems to agree with this hypothesis, as $\beta 5$ pro may be locked inside the 15S complex cavity by Ump1 so that it cannot access any complexes outside the cavity.

In summary, cross-linking and *in-vitro* binding assays enable a more thorough analysis of the functions of Ump1 and how the different parts of the protein relate to this. First of all the protein is unstructured so it is unlikely to provide a structural basis for the incorporation of various β subunits as previously assumed. Functionally speaking, the protein can be split into two parts: the N-terminal one (amino acid 1-50) which exits the 15S complex cavity via the β ring pore and interacts with other half-proteasomes, playing a part in their dimerization. The C-terminal one (amino acid 51-148) handles binding to the proteasome precursor complex and likely provides a checkpoint that prevents the dimerization of half-proteasomes until the final subunit has been incorporated into the β ring. These dual functions may provide an explanation to the seemingly paradoxical roles of Ump1 that have been reported before, being both a promoter and an inhibitor of 20S proteasome biogenesis.

7.3 An updated model of chaperone-assisted 20S proteasome biogenesis:

Using data gained from the analysis of the chaperones Pba1-Pba2 and Ump1, an updated model can be provided that traces the biogenesis of the 20S proteasome from the 15S precursor complex to the fully matured particle (figure 35). A detailed analysis of the state of the β rings in the 20S *pre1-1* complex is missing from the model, because a thorough biochemical analysis of this complex that could put the structural data into context is unavailable at the moment. For example the increased distance between the two β rings of the 20S *pre1-1* complex (which is also indicated in figure 35) has been observed before in inactive precursor complexes of archaeal 20S proteasomes (Groll et al. 2003). However, due to the lack of detailed biochemical information concerning all β subunit pro-peptides in the precursor complex from *S. cerevisiae* (not just those of catalytic subunits) the reason for this remains to be investigated. Nevertheless, structural data from both the 15S and 20S *pre1-1*

complexes shows that 20S proteasome biogenesis is a highly dynamic process, requiring both helper proteins and structural rearrangements for the successful assembly of the protein complex.

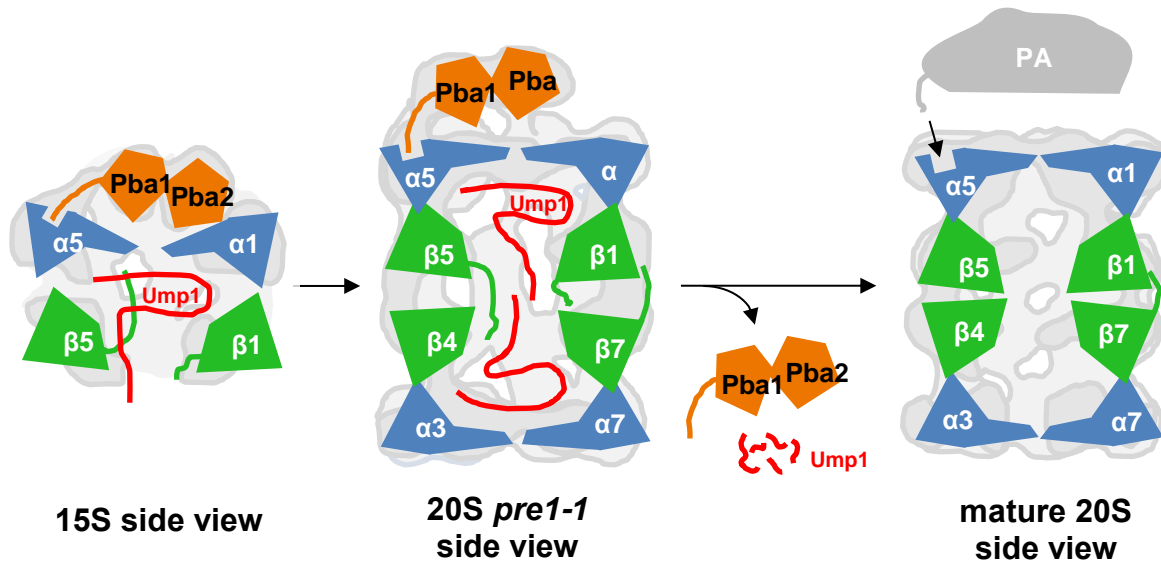


Figure 35: Structural model of proteasome biogenesis encompassing the stages of 15S to mature 20S. Schematics of the individual subunits are shown in the foreground with silhouettes of the respective EM reconstructions in the background. At the stage of the 15S precursor complex the chaperones Pba1-Pba2 are located in the middle of the pore, providing a structural scaffold for the binding of the α subunits through a deeper immersion in the α ring pore (essentially “plugging” the pore). The α subunits are pushed outward by the chaperones, which results in an overall larger diameter of the α ring (compared to the mature 20S) and an alteration of the pockets between the these subunits. The Pba1 and Pba2 HbYX motifs are unlikely to follow the canonical binding modes due to this alteration. However, the outside of the α ring is still blocked so that proteasome activators (PA) cannot bind. Ump1 is bound to the precursor complex via the subunits $\alpha 1$, $\alpha 4$, $\beta 4$, $\beta 6$ and the pro-peptide of $\beta 5$. Additional binding sites inside the cavity are possible. At this stage Ump1 keeps $\beta 5$ pro locked inside the complex. The β subunits are shifted into a conformation that is likely to inhibit dimerization as binding partners of two opposing β rings are not in the right positions. At the stage of maturation mimicked by the 20S *pre1-1* complex the half-proteasomes have dimerized. A contraction of the α ring has pushed Pba1-Pba2 out of the pore, which is now closed by the α subunit N-termini. The pockets between the α subunits adopt the same conformation that is also seen in the mature complex, meaning that the HbYX motifs of Pba1-Pba2 can adopt the canonical binding modes between the subunits $\alpha 5/\alpha 6$ and $\alpha 6/\alpha 7$. Since the α ring is fully matured but the rest of the proteasome is not, the main responsibility of the chaperone heterodimer appears to be blocking the access to activators. Incomplete maturation of the β ring (including pro-peptide cleavage) results in a partial separation of both halves of the proteasome, increasing the overall height of the complex compared to the mature particle. Which conformation Ump1 adapts at this stage is unknown. It is however likely that the incorporation of $\beta 7$ prior to half-proteasome dimerization has changed the Ump1 conformation to allow for a release of the $\beta 5$ pro-peptide. After the β subunit pro-peptides have been processed Ump1 is degraded by the newly activated subunits $\beta 1$, $\beta 2$ and $\beta 5$. The pro-peptide cleavage also allows all β subunits to adopt the conformation seen in the mature particle. Pba1-Pba2 dissociate from the mature 20S proteasome and are recycled, not degraded. The vacant HbYX motif binding pockets allow for binding of proteasome activators like Blm10 or the 19S RP. This concludes the proteasome biogenesis.

8. References:

- Bader, N. & Grune, T., 2006. Protein oxidation and proteolysis. *Biological Chemistry*, 387(10-11), pp.1351–5.
- Bochtler, M. et al., 1999. The proteasome. *Annual Review of Biophysics and Biomolecular Structure*, 28, pp.295–317.
- Burri, L. et al., 2000. Identification and characterization of a mammalian protein interacting with 20S proteasome precursors. *Proceedings of the National Academy of Sciences of the United States of America*, 97(19), pp.10348–10353.
- Chen, P. & Hochstrasser, M., 1996. Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell*, 86(6), pp.961–972.
- Chen, Q. et al., 2006. Ump1 extends yeast lifespan and enhances viability during oxidative stress: central role for the proteasome? *Free Radical Biology & Medicine*, 40(1), pp.120–126.
- Chesnel, F. et al., 2006. Cyclin B dissociation from CDK1 precedes its degradation upon MPF inactivation in mitotic extracts of *Xenopus laevis* embryos. *Cell Cycle*, 5(15), pp.1687–1698.
- Crowther, R.A., Henderson, R. & Smith, J.M., 1996. MRC image processing programs. *Journal of Structural Biology*, 16(4), pp.9–16.
- David, D.C. et al., 2002. Proteasomal degradation of tau protein. *Journal of Neurochemistry*, 83(1), pp.176–185.
- DeLano, W.L., 2002. The PyMOL molecular graphics system. *Schrödinger LLC*.
- Dick, T.P. et al., 1998. Contribution of proteasomal beta-subunits to the cleavage of peptide substrates analyzed with yeast mutants. *The Journal of Biological Chemistry*, 273(40), pp.25637–25646.
- Enenkel, C., Lehmann, A. & Klotzel, P.M., 1998. Subcellular distribution of proteasomes implicates a major location of protein degradation in the nuclear envelope-ER network in yeast. *The EMBO Journal*, 17(21), pp.6144–6154.
- Fehlker, M. et al., 2003. Bln3 is part of nascent proteasomes and is involved in a late stage of nuclear proteasome assembly. *EMBO Reports*, 4(10), pp.959–963.
- Frank, J. et al., 1996. SPIDER and WEB : Processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology*, 116(1), pp.190–199.
- Friant, S., Meier, K.D. & Riezman, H., 2003. Increased ubiquitin-dependent degradation can replace the essential requirement for heat shock protein induction. *The EMBO Journal*, 22(15), pp.3783–3791.

- Fricke, B. et al., 2007. The proteasome maturation protein POMP facilitates major steps of 20S proteasome formation at the endoplasmic reticulum. *EMBO Reports*, 8(12), pp.1170–1175.
- Fuller, S.D., 1987. The T=4 envelope of Sindbis virus is organized by interactions with a complementary T=3 capsid. *Cell*, 48(6), pp.923–934.
- Gerards, W.L. et al., 1997. The human alpha-type proteasomal subunit HsC8 forms a double ringlike structure, but does not assemble into proteasome-like particles with the beta-type subunits HsDelta or HsBPROS26. *The Journal of Biological Chemistry*, 272(15), pp.10080–10086.
- Grigorieff, N., 2007. FREALIGN: high-resolution refinement of single particle structures. *Journal of Structural Biology*, 157(1), pp.117–125.
- Groll, M. et al., 2000. A gated channel into the proteasome core particle. *Nature Structural Biology*, 7(11), pp.1062–1067.
- Groll, M., 2004. Habilitation: Strukturelle und funktionelle Zusammenhänge und Unterschiede archaebakterieller und eukaryontischer 20S-Proteasome.
- Groll, M. et al., 2003. Investigations on the maturation and regulation of archaebacterial proteasomes. *Journal of Molecular Biology*, 327(1), pp.75–83.
- Havens, C.G. et al., 2006. Regulation of late G1/S phase transition and APC Cdh1 by reactive oxygen species. *Molecular and Cellular Biology*, 26(12), pp.4701–4711.
- Van Heel, M., 1987. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, 21(2), pp.111–123.
- Van Heel, M., Harauz, G. & Orlova, E., 1996. A new generation of the IMAGIC image processing system. *Journal of Structural Biology*, 24(116), pp.17–24.
- Heinemeyer, W. et al., 1993. PRE2, highly homologous to the human major histocompatibility complex-linked RING10 gene, codes for a yeast proteasome subunit necessary for chymotryptic activity and degradation of ubiquitinated proteins. *The Journal of Biological Chemistry*, 268(7), pp.5115–5120.
- Heinemeyer, W. et al., 1997. The active sites of the eukaryotic 20S proteasome and their involvement in subunit precursor processing. *The Journal of Biological Chemistry*, 272(40), pp.25200–25209.
- Heinemeyer, W., Ramos, P.C. & Dohmen, R.J., 2004. The ultimate nanoscale mincer: assembly, structure and active sites of the 20S proteasome core. *Cellular and Molecular Life Sciences*, 61(13), pp.1562–1578.
- Heink, S. et al., 2005. IFN-gamma-induced immune adaptation of the proteasome system is an accelerated and transient response. *Proceedings of the National Academy of Sciences of the United States of America*, 102(26), pp.9241–9246.

- Henderson, R. et al., 2011. Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, 413(5), pp.1028–1046.
- Herzog, F. et al., 2012. Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science*, 337(6100), pp.1348–1352.
- Hirano, Y. et al., 2005. A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature*, 437(7063), pp.1381–1385.
- Hirano, Y. et al., 2006. Cooperation of multiple chaperones required for the assembly of mammalian 20S proteasomes. *Molecular Cell*, 24(6), pp.977–984.
- Jacquemont, C. & Taniguchi, T., 2007. Proteasome function is required for DNA damage response and fanconi anemia pathway activation. *Cancer Research*, 67(15), pp.7395–405.
- Jäger, S. et al., 1999. Proteasome beta-type subunits: unequal roles of propeptides in core particle maturation and a hierarchy of active site function. *Journal of Molecular Biology*, 291(4), pp.997–1013.
- Kleijnen, M.F. et al., 2007. Stability of the proteasome can be regulated allosterically through engagement of its proteolytic active sites. *Nature Structural & Molecular Biology*, 14(12), pp.1180–1188.
- Kusmierczyk, A.R. et al., 2011. A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nature Structural & Molecular Biology*, 18(5), pp.622–629.
- Kusmierczyk, A.R. et al., 2008. A multimeric assembly factor controls the formation of alternative 20S proteasomes. *Nature Structural & Molecular Biology*, 15(3), pp.237–244.
- Lehmann, A. et al., 2002. 20S proteasomes are imported as precursor complexes into the nucleus of yeast. *Journal of Molecular Biology*, 317(3), pp.401–413.
- Leitner, a. et al., 2012. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Molecular & Cellular Proteomics*, 11(3), p.M111.014126.
- Li, X. et al., 2007. beta-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. *The EMBO Journal*, 26(9), pp.2339–2349.
- Ludtke, S.J., Baldwin, P.R. & Chiu, W., 1999. EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology*, 128(1), pp.82–97.
- Marques, A.J. et al., 2007. The C-terminal extension of the beta7 subunit and activator complexes stabilize nascent 20 S proteasomes and promote their maturation. *The Journal of Biological Chemistry*, 282(48), pp.34869–34876.

- Mindell, J. & Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology*, 142(3), pp.334–347.
- Murata, S. et al., 2007. Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science*, 316(5829), pp.1349–1353.
- Orlowski, M. & Wilk, S., 2000. Catalytic activities of the 20 S proteasome, a multicatalytic proteinase complex. *Archives of Biochemistry and Biophysics*, 383(1), pp.1–16.
- Pettersen, E.F. et al., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), pp.1605–1612.
- Radermacher, M. et al., 1987. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *Journal of Microscopy*, 146(2), pp.113–136.
- Ramos, P.C. et al., 2004. Role of C-terminal extensions of subunits beta2 and beta7 in assembly and activity of eukaryotic proteasomes. *The Journal of Biological Chemistry*, 279(14), pp.14323–14330.
- Ramos, P.C. et al., 1998. Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell*, 92(4), pp.489–499.
- Ramos, P.C. & Dohmen, R.J., 2008. PACemakers of proteasome core particle assembly. *Structure*, 16(9), pp.1296–1304.
- Rape, M. & Jentsch, S., 2002. Taking a bite: proteasomal protein processing. *Nature Cell Biology*, 4(5), pp.E113–E116.
- Roseman, A.M., 2004. FindEM--a fast, efficient program for automatic selection of particles from electron micrographs. *Journal of Structural Biology*, 145(1-2), pp.91–9.
- Rosenthal, P.B. & Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, 333(4), pp.721–745.
- Saeki, Y. & Tanaka, K., 2007. Unlocking the proteasome door. *Molecular Cell*, 27(6), pp.865–867.
- Sá-Moura, B. et al., 2013. Biochemical and biophysical characterization of recombinant yeast proteasome maturation factor Ump1. *Computational and Structural Biotechnology Journal*, 7(8).
- Schmidtke, G. et al., 1996. Analysis of mammalian 20S proteasome biogenesis: the maturation of beta-subunits is an ordered two-step mechanism involving autocatalysis. *The EMBO Journal*, 15(24), pp.6887–6898.
- Shang, F. & Taylor, A., 2011. Ubiquitin-proteasome pathway and cellular responses to oxidative stress. *Free Radical Biology & Medicine*, 51(1), pp.5–16.

- Sharon, M. et al., 2006. 20S proteasomes have the potential to keep substrates in store for continual degradation. *The Journal of Biological Chemistry*, 281(14), pp.9569–9575.
- Sheaff, R.J. et al., 2000. Proteasomal turnover of p21Cip1 does not require p21Cip1 ubiquitination. *Molecular Cell*, 5(2), pp.403–410.
- Smith, D.M. et al., 2008. Docking of the proteasomal ATPases' C-termini in the 20S proteasomes alpha ring opens the gate for substrate entry. *Molecular Cell*, 27(5), pp.731–744.
- Sorzano, C.O.S. et al., 2004. XMIPP: a new generation of an open-source image processing package for electron microscopy. *Journal of Structural Biology*, 148(2), pp.194–204.
- Stadtmueller, B.M. et al., 2012. Structure of a Proteasome Pba1-Pba2 Complex: implications for proteasome assembly, activation, and biological function. *Journal of Biological Chemistry*, 287(44), pp.37371–37382.
- Takalo, M. et al., 2013. Protein aggregation and degradation mechanisms in neurodegenerative diseases. *American Journal of Neurodegenerative Diseases*, 2(1), pp.1–14.
- Le Tallec, B. et al., 2007. 20S proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and in mammals. *Molecular Cell*, 27(4), pp.660–674.
- Taylor, E.B. & Rutter, J., 2011. Mitochondrial quality control by the ubiquitin–proteasome system. *Biochemical Society Transactions*, 39, pp.1509–1513.
- Tofaris, G.K., Layfield, R. & Spillantini, M.G., 2001. alpha-Synuclein metabolism and aggregation is linked to ubiquitin-independent degradation by the proteasome. *FEBS Letters*, 509(1), pp.22–6.
- Uekusa, Y. et al., 2013. Backbone (1)H, (13)C, and (15)N assignments of yeast Ump1, an intrinsically disordered protein that functions as a proteasome assembly chaperone. *Biomolecular NMR Assignments*.
- Voges, D., Zwickl, P. & Baumeister, W., 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry*, 68, pp.1015–1068.
- Wang, J. & Maldonado, M. a, 2006. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. *Cellular & Molecular Immunology*, 3(4), pp.255–261.
- Yaffe, M.P. & Schatz, G., 1984. Two nuclear mutations that block mitochondrial protein import in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 81(15), pp.4819–4823.
- Yang, F., Moss, L.G. & Phillips, G.N., 1996. The molecular structure of green fluorescent protein. *Nature Biotechnology*, 14(10), pp.1246–1251.

- Yashiroda, H. et al., 2008. Crystal structure of a chaperone complex that contributes to the assembly of yeast 20S proteasomes. *Nature Structural & Molecular Biology*, 15(3), pp.228–236.

9. Appendix:

Appendix table 1: List of inter-subunit cross-links found for α and β subunits in 15S and 15S^{GFP}. Sub.1 and Sub.2 list the subunits, while Pos.1 and Pos.2 indicate the amino acid positions of the cross-linked lysines. The columns labelled 15S^{GFP} and 15S show in which complex the cross-link was found. Dist.[Å] lists the distance between C α atoms of the cross-linked lysines as per the subunit fit shown in figure 26. Distances labelled “?” indicate that measuring the distance was not possible, because one or both of the cross-linked lysines are situated in a protein or part of a protein for which no crystal structure is available. The first amino acid of all β subunits is always Thr1 or analogues residues as shown in figure 4. The preceding residue is labelled -1. Hence a negative number indicates pro-peptide or NTE residues.

Sub. 1	Sub. 2	Pos. 1	Pos. 2	15S ^{GFP}	15S	Dist. [Å]
$\alpha 1$	$\alpha 2$	167	50	x	x	24.8
$\alpha 1$	$\alpha 2$	33	50	x	x	28.7
$\alpha 1$	$\alpha 7$	58	178	x		19.1
$\alpha 1$	$\beta 1$	232	40	x	x	22
$\alpha 2$	$\alpha 3$	108	68	x	x	15.7
$\alpha 2$	$\alpha 3$	29	51	x		20.7
$\alpha 2$	$\alpha 3$	17	51	x		11.2
$\alpha 2$	$\beta 2$	98	33	x	x	24
$\alpha 2$	$\beta 3$	108	68	x	x	16.2
$\alpha 2$	$\beta 3$	98	183	x		17.8
$\alpha 2$	$\beta 3$	98	188	x		26.7
$\alpha 2$	$\alpha 3$	98	68		x	24.6
$\alpha 2$	$\alpha 3$	108	236		x	26
$\alpha 3$	$\alpha 2$	51	172	x	x	21.7
$\alpha 3$	$\alpha 2$	51	177	x		27.5
$\alpha 3$	$\beta 3$	231	66	x		27.9
$\alpha 3$	$\beta 3$	68	66	x	x	16.7
$\alpha 3$	$\beta 3$	100	61	x	x	13.6
$\alpha 3$	$\beta 4$	100	89	x		22.3
$\alpha 3$	$\beta 4$	72	112	x		18.6
$\alpha 4$	$\beta 4$	66	28	x		29.3
$\alpha 4$	$\beta 4$	66	108	x		28.9
$\alpha 4$	$\alpha 5$	182	241		x	29.9
$\alpha 4$	$\alpha 5$	182	52		x	21.4
$\alpha 5$	$\alpha 4$	211	182	x		22.2
$\alpha 5$	$\beta 5$	66	71		x	16.2
$\alpha 6$	$\alpha 7$	181	57	x	x	14.3
$\alpha 6$	$\alpha 7$	181	63	x	x	16.4
$\alpha 6$	$\beta 6$	65	74		x	24.6
$\alpha 7$	$\alpha 1$	173	58	x	x	17.1

Sub. 1	Sub. 2	Pos. 1	Pos. 2	15S ^{GFP}	15S	Dist. [Å]
$\alpha 7$	$\alpha 1$	100	98	x		20.3
$\alpha 7$	$\alpha 6$	57	169	x	x	7.9
$\alpha 7$	$\alpha 6$	208	169	x		21.4
$\alpha 7$	$\alpha 6$	63	102		x	17.3
$\beta 2$	$\alpha 1$	84	107	x	x	14.4
$\beta 2$	$\alpha 2$	29	98	x	x	27.2
$\beta 2$	$\beta 1$	84	33	x		21.1
$\beta 2$	$\beta 4$	-6	169		x	?
$\beta 3$	$\alpha 3$	183	68	x		21
$\beta 3$	$\beta 2$	109	204	x		22.8
$\beta 3$	$\beta 2$	109	199	x	x	26.4
$\beta 3$	$\beta 2$	109	29	x	x	21.1
$\beta 3$	$\beta 4$	30	185	x		24.9
$\beta 3$	$\beta 4$	30	112	x	x	17.5
$\beta 3$	$\alpha 3$	68	68		x	19.7
$\beta 4$	$\alpha 3$	112	218	x	x	20.6
$\beta 4$	$\alpha 3$	110	218	x	x	15.4
$\beta 4$	$\alpha 3$	109	218	x		14.6
$\beta 4$	$\alpha 3$	108	218	x	x	18.1
$\beta 4$	$\beta 3$	89	61	x	x	22.5
$\beta 4$	$\beta 3$	110	30	x	x	20.7
$\beta 4$	$\beta 3$	109	30		x	24.4
$\beta 5$	$\beta 4$	107	28		x	18.7
$\beta 5$	$\beta 4$	71	89		x	29.1
$\beta 5$	$\beta 4$	-7	28		x	?
$\beta 5$	$\alpha 6$	-60	115		x	?
$\beta 6$	$\alpha 6$	60	102	x	x	17.3
$\beta 6$	$\alpha 6$	64	102		x	13
$\beta 6$	$\alpha 6$	73	102		x	17.7

Appendix table 2: List of inter-subunit cross-links found between the chaperones Pba1, Pba2, Ump1 and regular proteasome subunits. Columns are labelled same as in appendix table 1.

Sub.1	Sub.2	Pos.1	Pos.2	15S ^{GFP}	15S	Dist.[Å]
Pba1	$\alpha 3$	91	51		x	?
Pba2	$\alpha 7$	261	57	x	x	21
Pba2	$\alpha 7$	261	167	x	x	19.6
Pba2	Pba1	65	213		x	13.1
Ump1	$\alpha 1$	83	98	x		?
Ump1	$\alpha 1$	113	107	x		?
Ump1	$\alpha 4$	144	88	x	x	?
Ump1	$\beta 4$	144	89	x	x	?
Ump1	$\beta 6$	19	91		x	?

Appendix table 3: List of inter-subunit cross-links found between α and β subunits that are not consistent with the subunit fit shown in figure 26 or where the cross-link distance exceeds 30 Å. Columns are labelled same as in appendix table 1.

Sub. 1	Sub. 2	Pos. 1	Pos. 2	15S ^{GFP}	15S	Dist. [Å]
$\alpha 1$	$\alpha 6$	167	61	x		80
$\alpha 1$	$\alpha 6$	119	61	x		69.6
$\alpha 4$	$\alpha 1$	193	98	x		89
$\alpha 4$	$\alpha 3$	39	68	x		57.5
$\alpha 6$	$\beta 2$	65	148		x	92.3
$\alpha 6$	$\beta 4$	65	169		x	82
$\alpha 7$	$\beta 2$	100	199	x		78.3
$\alpha 7$	$\beta 2$	100	201	x		82.6
$\alpha 7$	$\beta 3$	208	183		x	97.1
$\beta 2$	$\alpha 4$	199	169	x		98.7
$\beta 4$	$\alpha 6$	169	102		x	80.2
$\beta 5$	$\alpha 6$	71	102		x	55

Appendix table 4: List of inter-subunit cross links found between the GFP moiety fused to Ump1 in the 15S^{GFP} complex and other proteasome subunits. Columns are labelled same as in appendix table 1. Distances are not shown because an unambiguous fit for the GFP crystal structure could not be established.

Sub.1	Sub.2	Pos.1	Pos.2
GFP	$\alpha 1$	41	98
GFP	$\alpha 1$	101	98
GFP	$\alpha 1$	166	98
GFP	$\alpha 2$	107	98
GFP	$\alpha 4$	79	88
GFP	$\alpha 6$	101	102
GFP	$\alpha 6$	41	115
GFP	$\alpha 6$	107	102
GFP	$\alpha 6$	41	61
GFP	$\alpha 6$	79	102
GFP	$\alpha 6$	156	61
GFP	$\alpha 6$	156	115
GFP	$\alpha 6$	166	102
GFP	$\alpha 6$	79	115

Sub.1	Sub.2	Pos.1	Pos.2
GFP	$\alpha 6$	101	167
GFP	$\alpha 7$	101	63
GFP	$\alpha 7$	41	100
GFP	$\alpha 7$	166	100
GFP	$\alpha 7$	166	104
GFP	$\alpha 7$	79	100
GFP	$\alpha 7$	101	104
GFP	$\beta 2$	166	84
GFP	$\beta 2$	107	199
GFP	$\beta 4$	41	28
GFP	$\beta 4$	156	28
GFP	$\beta 4$	156	89
GFP	$\beta 4$	79	89
GFP	Ump1	156	58

Appendix table 5: List of intra-subunit cross-links found for subunits in 15S and 15S^{GFP}. Sub. signifies the subunits, while Pos.1 and Pos.2 indicate the amino acid positions of the cross linked lysines. The columns labelled 15S^{GFP} and 15S show in which complex the cross-link was found.

Sub.	Pos.1	Pos.2	15S ^{GFP}	15S
$\alpha 1$	232	98	x	x
$\alpha 1$	190	48	x	x
$\alpha 1$	174	58	x	
$\alpha 1$	230	48	x	x
$\alpha 1$	167	48	x	x
$\alpha 1$	167	190	x	
$\alpha 1$	119	98	x	x
$\alpha 1$	48	98	x	x
$\alpha 1$	187	190	x	x
$\alpha 1$	33	174	x	x
$\alpha 1$	188	48	x	
$\alpha 1$	187	58	x	
$\alpha 2$	91	88	x	x
$\alpha 2$	50	166	x	x
$\alpha 2$	29	172	x	x
$\alpha 2$	237	49	x	x
$\alpha 2$	29	166	x	x
$\alpha 2$	98	88	x	x
$\alpha 2$	29	177	x	
$\alpha 2$	166	49	x	
$\alpha 2$	237	50		x
$\alpha 3$	231	68	x	x
$\alpha 3$	228	241	x	x
$\alpha 3$	247	236	x	x
$\alpha 3$	247	241	x	x
$\alpha 3$	228	68	x	x
$\alpha 3$	247	199	x	x
$\alpha 3$	247	195	x	
$\alpha 3$	247	231	x	x
$\alpha 3$	246	199	x	x
$\alpha 3$	247	218	x	
$\alpha 3$	185	218	x	
$\alpha 3$	246	236	x	x
$\alpha 3$	181	218	x	
$\alpha 3$	181	195	x	
$\alpha 3$	236	68		x
$\alpha 3$	231	246		x
$\alpha 3$	100	68		x
$\alpha 4$	53	63	x	x
$\alpha 4$	182	53	x	

Sub.	Pos.1	Pos.2	15S ^{GFP}	15S
$\alpha 4$	182	37	x	x
$\alpha 4$	39	63	x	x
$\alpha 4$	182	39	x	x
$\alpha 4$	53	169	x	x
$\alpha 4$	37	53	x	x
$\alpha 4$	248	169	x	x
$\alpha 4$	177	169	x	x
$\alpha 4$	53	251	x	
$\alpha 4$	182	169	x	x
$\alpha 4$	39	53	x	x
$\alpha 4$	53	248	x	
$\alpha 4$	39	248	x	
$\alpha 4$	182	248	x	x
$\alpha 4$	177	248	x	x
$\alpha 4$	182	177	x	x
$\alpha 4$	177	250	x	
$\alpha 4$	182	249	x	
$\alpha 4$	177	251	x	
$\alpha 4$	208	53	x	
$\alpha 4$	63	88		x
$\alpha 4$	28	169		x
$\alpha 5$	43	224	x	x
$\alpha 5$	211	241	x	x
$\alpha 5$	246	205	x	x
$\alpha 6$	65	102	x	x
$\alpha 6$	50	61	x	x
$\alpha 6$	30	169	x	
$\alpha 6$	65	61	x	x
$\alpha 6$	115	102		x
$\alpha 6$	169	102		x
$\alpha 6$	61	102		x
$\alpha 6$	50	102		x
$\alpha 7$	66	167	x	x
$\alpha 7$	194	167	x	x
$\alpha 7$	57	66	x	
$\alpha 7$	63	167	x	
$\alpha 7$	208	167	x	x
$\alpha 7$	178	194	x	x
$\alpha 7$	208	66	x	
$\alpha 7$	232	63	x	x

Appendix table 5 (continued).

Sub.	Pos.1	Pos.2	15S ^{GFP}	15S
$\alpha 7$	208	57	x	x
$\alpha 7$	52	63	x	x
$\alpha 7$	178	198	x	x
$\alpha 7$	173	194	x	x
$\alpha 7$	173	167	x	x
$\alpha 7$	57	167	x	
$\alpha 7$	29	167	x	
$\alpha 7$	178	167	x	x
$\alpha 7$	208	194	x	x
$\alpha 7$	178	208		x
$\alpha 7$	232	194		x
$\alpha 7$	29	173		x
$\alpha 7$	104	100		x
$\beta 1$	40	33	x	x
$\beta 1$	107	40	x	
$\beta 1$	140	148	x	
$\beta 2$	204	199	x	x
$\beta 2$	29	204	x	x
$\beta 2$	29	201	x	
$\beta 2$	29	199	x	x
$\beta 2$	-6	29	x	x
$\beta 2$	-6	199	x	
$\beta 2$	33	199	x	
$\beta 3$	30	192	x	x
$\beta 3$	188	192	x	x
$\beta 3$	61	68	x	x
$\beta 3$	183	192	x	x
$\beta 3$	30	188	x	x
$\beta 3$	66	183	x	
$\beta 3$	109	183	x	
$\beta 3$	30	183		x
$\beta 4$	28	181	x	x
$\beta 4$	108	185	x	x
$\beta 4$	28	176	x	x
$\beta 4$	18	28	x	x
$\beta 4$	176	33	x	x
$\beta 4$	18	181	x	x
$\beta 4$	181	33	x	x
$\beta 4$	176	169	x	x
$\beta 4$	18	176	x	x

Sub.	Pos.1	Pos.2	15S ^{GFP}	15S
$\beta 4$	112	185	x	x
$\beta 4$	18	33	x	
$\beta 4$	110	185	x	
$\beta 4$	185	28	x	
$\beta 4$	109	185	x	x
$\beta 4$	28	176		x
$\beta 4$	18	176		x
$\beta 5$	107	71		x
$\beta 5$	-5	71		x
$\beta 5$	32	-5		x
$\beta 6$	205	60	x	x
$\beta 6$	200	74		x
Pba1	35	81	x	
Pba1	23	81		x
Ump1	58	144	x	
GFP	101	166	x	
GFP	126	107	x	
GFP	131	101	x	
GFP	27	19	x	
GFP	107	162	x	
GFP	107	126	x	
GFP	101	162	x	
GFP	126	101	x	
GFP	107	158	x	
GFP	113	126	x	

Appendix table 6: List of mono cross-links found for subunits in 15S and 15S^{GFP}. Sub. signifies the subunit, while Pos. indicates the amino acid position of the cross-linked lysine. The columns labelled 15S^{GFP} and 15S show in which complex the cross link was found.

Sub.	Pos.	15S ^{GFP}	15S
$\alpha 1$	33	x	x
$\alpha 1$	48	x	x
$\alpha 1$	58	x	x
$\alpha 1$	62	x	x
$\alpha 1$	107	x	x
$\alpha 1$	167	x	x
$\alpha 1$	174	x	x
$\alpha 1$	187	x	x
$\alpha 1$	188	x	x
$\alpha 1$	190	x	x
$\alpha 1$	232	x	x
$\alpha 2$	17	x	
$\alpha 2$	29	x	x
$\alpha 2$	50	x	
$\alpha 2$	63	x	
$\alpha 2$	98	x	x
$\alpha 2$	108	x	x
$\alpha 2$	237	x	
$\alpha 3$	51	x	
$\alpha 3$	65	x	x
$\alpha 3$	100	x	x
$\alpha 3$	195	x	x
$\alpha 3$	199	x	x
$\alpha 3$	228	x	x
$\alpha 3$	231	x	x
$\alpha 3$	236	x	x
$\alpha 3$	241	x	x
$\alpha 3$	246	x	x
$\alpha 3$	247	x	x
$\alpha 3$	258	x	x
$\alpha 4$	28	x	
$\alpha 4$	39		x
$\alpha 4$	53	x	x
$\alpha 4$	66	x	x
$\alpha 4$	88	x	x
$\alpha 4$	182	x	x
$\alpha 4$	193	x	x
$\alpha 5$	43	x	x
$\alpha 5$	66	x	x
$\alpha 5$	170		x

Sub.	Pos.	15S ^{GFP}	15S
$\alpha 5$	194	x	
$\alpha 5$	202	x	
$\alpha 5$	205	x	x
$\alpha 5$	211	x	x
$\alpha 6$	50	x	x
$\alpha 6$	61	x	x
$\alpha 6$	62	x	
$\alpha 6$	65	x	x
$\alpha 6$	102		x
$\alpha 6$	115	x	x
$\alpha 6$	118	x	x
$\alpha 6$	139	x	
$\alpha 6$	181	x	
$\alpha 6$	191	x	x
$\alpha 6$	218	x	
$\alpha 6$	232	x	
$\alpha 7$	29	x	x
$\alpha 7$	52		x
$\alpha 7$	101	x	x
$\alpha 7$	104	x	x
$\alpha 7$	143	x	
$\alpha 7$	173	x	x
$\alpha 7$	178	x	x
$\alpha 7$	194	x	x
$\alpha 7$	198	x	x
$\alpha 7$	208	x	x
$\alpha 7$	232	x	x
$\beta 1$	33	x	x
$\beta 1$	107	x	
$\beta 1$	121	x	
$\beta 1$	140	x	
$\beta 1$	156	x	x
$\beta 1$	164	x	x
$\beta 2$	-6	x	x
$\beta 2$	7	x	x
$\beta 2$	40	x	
$\beta 2$	84	x	x
$\beta 2$	89	x	
$\beta 2$	199	x	x
$\beta 3$	30	x	x

Appendix table 6 (continued).

Sub.	Pos.	15S ^{GFP}	15S
β3	109	x	
β4	18	x	x
β4	28	x	x
β4	33	x	
β4	108	x	x
β4	109	x	x
β4	110	x	x
β4	112	x	x
β4	124	x	
β4	176	x	x
β4	185	x	
β5	-60		x
β5	-7	x	x
β5	-5		x
β5	71	x	x
β5	81	x	x
β5	107		x
β6	158		x
β6	160	x	
β6	200		x
Pba1	13		x
Pba1	23	x	x
Pba1	35	x	
Pba1	63	x	x
Pba1	91	x	x
Pba1	213		x
Pba1	215	x	x
Pba2	60		x
Pba2	76		x
Pba2	77		x
Pba2	187	x	
Pba2	200	x	x
Pba2	232	x	
Pba2	235	x	
Pba2	242	x	x
Ump1	19		x
Ump1	58	x	x
Ump1	141	x	
Ump1	144	x	x
GFP	41	x	

Sub.	Pos.	15S ^{GFP}	15S
GFP	85	x	
GFP	107	x	
GFP	113	x	
GFP	126	x	
GFP	131	x	
GFP	156	x	

10. Abbreviations:

APS	Ammonium persulfate
ATP	Adenosine triphosphate
β 5pro	Pro-peptide of the β 5 subunit (analogous for β 1pro and β 2pro)
BSA	Bovine serum albumin
CCD	Charge-coupled device
ChyT-I	Chymotrypsin-like
CP	Core particle
CTE	C-terminal extension
CTF	Contrast transfer function
DMF	N,N-Dimethylformamide
DNA	Desoxy-ribonucleic acid
DSS	Disuccinimidyl suberate
DTT	Dithiothreitol
e ⁻	Electron
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EtOH	Ethanol
FSC	Fourier shell correlation
GFP	Green fluorescent protein
HA	Human influenza hemagglutinin
HAc	Acetic acid
HbYX	Hydrophobic amino acid followed by tyrosine and a random amino acid
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
his	Histidin
HRP	Horseradish peroxidase
IgG	Immunoglobulin protein G
IPTG	Isopropyl- β -D-thiogalactopyranosid
kDa	Kilodalton
LB	Luria broth
leu	Leucin
MAT	Mating type

MeOH	Methanol
MS	Mass spectrometry
MSA	Multivariate statistical analysis
MRA	Multi reference alignment
NaDOC	Sodium deoxycholate
NHS	N-hydroxysuccinimide
NTA	Nitrilotriacetic acid
NTE	N-terminal extension
OD ₆₀₀	Optical density at 600 nanometer wavelength
RP	Regulatory particle
rpm	Revolutions per minute
PA	Proteasome activator
PAC	Proteasome assembling chaperone
PAGE	Polyacrylamide gel electrophoresis
Pba	Proteasome biogenesis associated
PGPH	Peptidyl-glutamyl peptide-hydrolyzing
ProA	Protein A
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Sodium dodecyl sulfate
siRNA	Small interfering ribonucleic acid
T-I	Trypsin-like
TBE	Tris, boric acid, EDTA
TCA	Trichloroacetic acid
TEM	Transmission electron microscopy
TEMED	N,N,N',N'-Tetramethylethylenediamine
TEV	Tobacco etch virus
Tris	Tris(hydroxymethyl)aminomethane
TST	Tris buffered saline, Tween-20
UA	Uranyl acetate
Ump1	Ubiquitin mediated proteolysis 1
ura	Uracil
YPD	Yeast peptone D-glucose

11. Acknowledgements:

First and foremost I would like to thank Petra Wendler for supervising my thesis and teaching me all about proteasomes, electron microscopy and how to process the data. I would also like to thank my thesis advisory committee members Roland Beckmann and Kristina Lakomek for helpful advice during the years I spent on my thesis.

Additional thanks go to Jürgen Dohmen and Paula Ramos for the ideas that got this project started, as well as fruitful discussions about proteasome precursor complexes. Special thanks go to Maria Nunes for providing me with some truly excellent protein samples for electron microscopy and MS. Franz Herzog and Matthias Hemann deserve a shout-out for providing me with MS data that enabled an independent verification of my findings.

I would also like to thank Sebastian Kube and Susanne Ciniawsky for all the help concerning EM sample preparation, data processing, etc. and the occasional encouragement during frustrating times in the lab.